LEIPZIG

Band 74

Karl-Heinz Eger

Sequential Tests

TEUBNER-TEXTE zur Mathematik

TEUBNER-TEXT

Dr. Karl-Heinz Eger

Born in 1945 in Penig. Studied Mathematics in Karl-Marx-Stadt. Since 1968 at the Department of Mathematics at the Karl-Marx-Stadt Technische Hochschule. Attendance at the term on 'Sequential Statistics' at Banach Center Warsaw in 1981.

Field of interest: Sequential statistics.

Eger, Karl-Heinz:
Sequential Tests. - Leipzig: BSB Teubner Verlagsges. 1985.
172 S. - (Teubner-Texte zur Mathematik; Bd. 74)

ISSN 0138-502X

1. Auflage
VLN 294-375/52/85 · LSV 1075
Lektor: Dr. Renate Müller
nted in the German Democratic Republic
Kongreß- und Werbedruck, Oberlungwitz
i: Messedruck Leipzig, Betriebsteil Taucha
666 209 3

TEUBNER-TEXTE zur Mathematik · Band 74

Karl-Heinz Eger

Sequential Tests

The present text pr[illegible] a systematic introduction to sequential tests. By means of a new conjugacy principle the properties of sequential likelihood ratio tests are investigated. Broad space is devoted to Wald's likelihood ratio test where constructive aspects like the design of admissible tests or the computation of the characteristics play an important part. A direct method of the computation of the characteristics of Wald's sequential likelihood ratio test is developed by which the computation of such characteristics like the power function or the moments of the sample size may be reduced to that of solving systems of linear equations. Moreover, a continuous inspection scheme based on sequential tests and the discrimination among more than two hypotheses are considered. A test for a simultaneous observation of several Bernoulli distributed random variables is discussed.

Der vorliegende Text ist eine systematische Einführung in die Theorie der sequentiellen Tests. Mit Hilfe eines neuen Konjugiertheitsprinzips werden die Eigenschaften des sequentiellen Quotiententests untersucht. Breiter Raum wird dem Waldschen sequentiellen Quotiententest eingeräumt, wobei konstruktive Gesichtspunkte wie die Konstruktion zulässiger Tests oder die Berechnung der Charakteristiken sequentieller Tests im Vordergrund stehen. Es wird eine direkte Methode zur Berechnung der Charakteristiken des Waldschen sequentiellen Quotiententests entwickelt, mit der die Berechnung solcher Charakteristiken, wie der Machtfunktion oder der Momente des Stichprobenumfangs, auf die Lösung linearer Gleichungssysteme zurückgeführt werden kann. Darüber hinaus wird ein kontinuierliches Stichprobenverfahren auf der Basis sequentieller Tests und die Unterscheidung zwischen mehr als zwei Hypothesen mit Hilfe sequentieller Tests betrachtet. Außerdem wird ein Test für die gleichzeitige Beobachtung mehrerer 0-1-verteilter Zufallsgrößen vorgestellt.

Le texte présent est une introduction systématique dans la théorie des tests séquentiels. À l'aide d'un nouveau principe de conjugaison on étude les propriétés du ratio-test séquentiel. On donne une attention particulière au ratio-test séquentiel de Wald, en considérant en premier lieu des aspects constructifs, comme la construction de tests admissibles ou la computation des caractéristiques de tests séquentiels. On développe une méthode directe pour la computation des caractéristiques du ratio-test séquentiel de Wald qui permet de ramener la computation de telles caractéristiques, comme de la fonction puissance ou des moments de la taille d'échantillons, au cas de la résolution de systèmes linéaires d'equations. De plus, on considère une méthode d'échantillons continue sur la base de tests séquentiels et la distinction entre plus de 2 hypothèses à l'aide tests séquentiels. En outre, on présente un test pour l'observation de plusieurs variables aléatoires avec une distribution à deux points en même temps.

Настоящая работа представляет собой систематическое введение в теорию последовательных критериев.
При помощи нового принципа сопряженности исследуются свойства последовательного критерия правдоподобия. Большое внимание уделяется последовательному критерию правдоподобия Вальда, причем в центре внимания стоят конструктивные аспекты, как, например, конструкция допустимых критериев или вычисление характеристик последовательных критериев.
Разработается прямой метод вычисления характеристик последовательного критерия правдоподобия Вальда, позволяющий свести вычисление таких характеристик, как функции мощности или моментов объема выборки, к решению систем линейных уравнений. Далее, рассматриваются непрерывный метод выборки на основе последовательных критериев и различение между более чем 2 чипотезами с помощью последовательных критериев. Кроме того, представляется критерий для одновременного наблюдения нескольких случайных величин с двухточечным распределением.

Preface

The purpose of this text is to provide readers with a systematic introduction to sequential tests. Since the publication of Abraham Wald's monograph 'Sequential Analysis' in the late forties theoretical and practical interest in this topic has grown rapidly. The increasing relevance of this field has also been effected by progress in computer technology. Present-day computer technology allows the uncomplicated realization of sequential tests and the effective computation of their characteristics. This is a significant aspect in respect of the design, assessment, and application of sequential tests.

Full exploitation of the advantages of sequential tests requires a corresponding theoretical foundation. Without such a foundation the effective design of sequential tests is difficult. Therefore, besides the explanation of the basic elements of the sequential test theory one aim of this text is to elaborate those elements of the theory which may be relevant to the design and application of sequential tests.

Chapter 1 serves to introduce the theory of sequential tests and their embedding in the general theory of sequential statistics. Those parts of the general theory are elaborated which are relevant for an investigation of the properties of sequential tests. In Section 1.6 a conjugacy principle is developed which will serve as a useful device in investigating the quantitative and qualitative properties of sequential tests. The reader who is more interested in design of sequential tests may consider Sections 1.1 to 1.5 more from an informatory point of view, whereas the conjugacy concept of Section 1.6 will be needed for a basic understanding of the subsequent sections.

Chapter 2 is devoted to the systematic investigation of the properties of sequential likelihood ratio tests in which, of course, Wald's likelihood ratio test plays a particular role. Besides dealing with the general properties of sequential likelihood ratio tests, constructive aspects will play an important part, for instance, the design of admissible tests or the approximate computation of the characteristics of Wald's likelihood ratio test.

Chapter 3 presents a method for the computation of characteristics

like the power function, the moments of the sample size, etc. of Wald's likelihood ratio test. By this method, originally developed for tests based on sequences of integer-valued random variables, the computation of the characteristics mentioned above can be reduced to that of solving systems of linear equations. Since many continuous test problems can be correspondingly discretized, this approach will also be of interest for continuous test problems. The method presented only needs vector and matrix operations and can be easily carried out on a computer. For this reason, the author is convinced that this method will be one method to be taken into consideration in the design of sequential tests in the future. Some new aspects of the design of sequential likelihood ratio tests are considered in Section 3.8. Moreover, a continuous inspection scheme based on sequential tests is investigated.

Finally, Chapters 4 and 5 are devoted to special problems, the discrimination among more than two hypotheses and the simultaneous observation of several Bernoulli distributed random variables.

It has not been possible within the scope of this little book to include extensive numerical examples, although the author has carried out detailed numerical investigations in connection with the procedures presented in this text. The author would welcome correspondence on this.

My interest in the present undertaking was aroused by Professor Heckendorff, to whom I wish to express my sincere thanks for his constant interest and encouragement. My thanks are due also to the TEUBNER publishing house for their excellent collaboration.

Hints of any kind will be appreciated by the author.

Karl-Marx-Stadt
January 1985

K.-H. Eger

Contents

1. General theory

The investigations carried out by A. WALD in the early fourties, concerning his famous sequential likelihood ratio test, may be considered as the origin of a new branch of mathematical statistics, the so-called sequential statistics. The beginning of this area is characterized by very special investigations in connection with sequential tests and certain estimation problems, and the corresponding results concerning sequential tests are contained in A. WALD's celebrated monograph 'Sequential Analysis' [77]. In the meantime, this subject has grown considerabely (see e.g. [46], [48]). Many results of this field are summarized in the monographs by GHOSH [35] and GOVINDARAJULU [37] . Parallel to the development of special sequential procedures, investigations have been carried out, concerning the structure of sequentiel statistical procedures. The investigations carried out by BAHADUR [8] are an essential contribution in this direction where the sufficiency principle is extended to sequential statistical structures. The theory of optimal stopping, which is relevant for sequential statistics, is elaborated in the books by CHOW et al. [22] and ŠIRJAJEV [73] . The book by HECKENDORFF [40] may be considered as a concentrated elaboration of the foundations of sequential statistical structures. A detailed explanation of the theory of sequential tests and the investigation into the properties of these tests is not possible any longer without reference to the general theory of sequential statistical structures. Otherwise, the space of this booklet does not allow to go too far into the details of the general theory. For this reason, we shall frequently refer to HECKENDORFF [40] in this context.

This first section is designed to give an introduction into the terminology of the sequential tests and their embedding into the general theory of sequential statistical structures. In Sections 1.1 to 1.3 the corresponding notations of the test theory are introduced. In Section 1.4 we shall consider some sufficiency properties of sequential statistical structures, which will play a part in sequential tests. The importance of the sufficiency concept for simplifying the structure of a sequential test by corresponding data reduction is discussed as well as the significance of sufficiency is emphasized in connection with the computation of the characteristics of a sequential test. The important role of the sequence of the likelihood ratios for sequential tests is elaborated, and certain convergence properties of this sequence are investigated in Section 1.5.

In Section 1.6, a new conjugacy approach is introduced which will serve as a useful tool in investigating the properties of sequential tests.

The reader, who is more interested in the practical design of sequential tests, may consider Sections 1.1 to 1.5 more from an informatory point of view, whereas the results of Section 1.6 will fundamentally be needed in the subsequent sections.

1.1 Introduction

Let $(\Omega, \mathfrak{F}, \mathfrak{P})$ be a statistical structure to any given experiment where $(\Omega, \mathfrak{F})$ denotes a measurable space and $\mathfrak{P}$ a family of possible probability measures on $(\Omega, \mathfrak{F})$. We suppose that the experiment can be carried out in successive steps or stages and that a corresponding sequence $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ of non-decreasing sub-σ-algebras of $\mathfrak{F}$ is given, completed by $\mathfrak{F}_\infty = \sigma(\bigcup_{n\in\Gamma^+} \mathfrak{F}_n)$, so that for every $n \in \Gamma^+$ the sub-σ- algebra $\mathfrak{F}_n$ can be interpreted as the set of all events which are observational until the n^{th} stage of the experiment. If we denote by $\mathfrak{P}_n$ the restriction of the family $\mathfrak{P}$ from $(\Omega, \mathfrak{F})$ to $(\Omega, \mathfrak{F}_n)$, we obtain a statistical structure $(\Omega, \mathfrak{F}_n, \mathfrak{P}_n)$ which can be regarded as a statistical structure for the first n stages of the experiment, or in a more general sense, as a statistical structure at the fixed sample size n, $n \in \Gamma^+$.

This approach can be generalized as follows. Let $N: \Omega \to \Gamma^+$ be a stopping time[1)] with respect to (w.r.t.) the sequence of σ-algebras $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$, then we may write this stopping time always as

$$N = \begin{cases} \inf\{n \geq 1: \chi_{\{N=n\}} = 1\} & \text{, if such an n exists,} \\ \infty & \text{, otherwise .} \end{cases} \tag{1.1}$$

Based on this stopping time N, we can obtain a random sample size in the following manner. Beginning with $n = 1$, we continue the experiment or continue sampling as long as $\chi_{\{N=n\}} = 0$ for $n = 1,2,...$ We stop the experiment or sampling after n stages if we observe $\chi_{\{N=n\}} = 1$ at the n^{th} sampling stage for the first time.

Each stopping time N w.r.t. $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ generates a σ-algebra $\mathfrak{F}_N$, defined by $\mathfrak{F}_N = \sigma\{F_n \cap \{N=n\}, F_n \in \mathfrak{F}_n, n \in \bar{\Gamma}^+\}$.

1) A measurable function $N: \Omega \to \Gamma^+$ is called a stopping time w.r.t. $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ if $\{N=n\} \in \mathfrak{F}_n$, $n \in \Gamma^+$.

This is the so-called σ-algebra of the N-past and it contains all those events of $\mathcal{F}$ which can be observed, using a sample size given by stopping time N. Thus, we obtain a corresponding statistical structure, say $(\Omega, \mathcal{F}_N, \mathcal{P}_N)$, where $\mathcal{P}_N$ denotes the restriction of the family $\mathcal{P}$ from $(\Omega, \mathcal{F})$ to $(\Omega, \mathcal{F}_N)$. Such a structure is called a sequential statistical structure. For further statistical details of such structures, we refer to HECKENDORFF [40] .

For the following investigations, it will be convenient to suppose that family $\mathcal{P}$ is indexed by a parameter θ that belongs to a parameter set Θ , so that $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ and card $\Theta \geq 2$. The statistical procedures under consideration are then defined as follows.

D e f i n i t i o n 1.1.1. Let $\{\mathcal{F}_n\}_{n \in \Gamma^+}$ be a non-decreasing sequence of sub-σ-algebras of $\mathcal{F}$ and N a stopping time w.r.t. $\{\mathcal{F}_n\}_{n \in \Gamma^+}$, let $\mathcal{F}_N$ be the corresponding σ-algebra of the N-past, and let $\delta: \Omega \to [0,1]$ be an $(\mathcal{F}_N, \mathcal{B}^1)$-measurable function. Then, to any given non-empty disjoint parameter sets $\Theta_0 \subset \Theta$ and $\Theta_1 \subset \Theta$, the pair (N, δ) is said to be a test with the sample size N and the terminal decision rule δ for the (null) hypothesis $H_0: \theta \in \Theta_0$ against the (alternative) hypothesis $H_1: \theta \in \Theta_1$ if we proceed as follows:

(i) We continue sampling as long as $\chi_{\{N=n\}} = 0$ holds for $n = 1,2,\ldots$ and stop sampling at the first stage n, $n = 1,2,\ldots$, where we observe $\chi_{\{N=n\}} = 1$.

(ii) If we stop sampling at stage $N(\omega) = n$, $n \in \Gamma^+$, we reject the hypothesis H_0 (accept H_1) with the probability $\delta(\omega)$, or accept the hypothesis H_0, respectively, if H_0 is not rejected by it. Briefly we shall say (N, δ) is a test based on $\{\mathcal{F}_n\}_{n \in \Gamma^+}$.

For a more detailed characterization of the properties of sample size N of any given test (N, δ) we introduce the following notations. A test (N, δ) is said to be a fixed-sample test, based on n observations, if an $n \in \Gamma^+$ exists, so that $N(\omega) = n$ for every $\omega \in \Omega$. Each other test is called a sequential test. A test (N, δ) is said to be closed iff for every $\theta \in \Theta$

$$P_\theta(N < \infty) = 1 . \tag{1.2}$$

(N, δ) is said to be open iff a $\theta' \in \Theta$ exists with $P_{\theta'}(N < \infty) < 1$. If a test (N, δ) is closed it is also usual to say the termination property holds or N is closed. A test (N, δ) is said to be truncated at stage $\bar{n}$, $\bar{n} \in \Gamma^+$, iff $\bar{n}$ is the smallest integer so that

for every $n \geq \bar{n}$ and $\theta \in \Theta$ $P_\theta(N \leq n) = 1$ holds.
With respect to the terminal decision rule, a test (N, δ) is said to be randomized if an $\omega \in \Omega$ exists, where $0 < \delta(\omega) < 1$, otherwise δ is said to be non-randomized. We refer to HECKENDORFF [40] , 1.6, for a more detailed description of randomized terminal decision rules. The hypothesis H_0: $\theta \in \Theta_0$ is said to be a simple hypothesis if Θ_0 contains only a single parameter, otherwise H_0 is said to be a composite hypothesis. In an analogous manner we distinguish between a simple and a composite alternative hypothesis H_1: $\theta \in \Theta_1$.
According to the definition of the test (N, δ), its properties are completely determined by the sample size N and the terminal decision rule δ. Let now $w: R^2 \rightarrow R^1$ be any given measurable function, then $w(N, \delta)$ is an $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable function. If the expectation value $E_\theta w(N, \delta)$ exists for every $\theta \in \Theta$, this expectation value as a function of θ reflects, depending on the choice of this function w, certain statistical properties of the test (N, δ). Therefore, we define the following.

D e f i n i t i o n 1.1.2. Let (N, δ) be any given test, and let $w: \Omega \rightarrow R^1$ be an $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable function where $E_\theta w$ exists for every $\theta \in \Theta$. Then the expectation value $E_\theta w$ as a function of $\theta \in \Theta$ is called a characteristic of the test (N, δ).

We remark that a function $w: \Omega \rightarrow R^1$ is $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable iff a sequence $\{w_n\}_{n \in \bar{\Gamma}^+}$ of $(\mathfrak{F}_n, \mathfrak{B}^1)$-measurable functions $w_n: \Omega \rightarrow R^1$ exists so that

$$w = \sum_{n \in \bar{\Gamma}^+} w_n \chi_{\{N = n\}}$$

holds (see [40] , Theorem 1.5). That means that we have also

$$E_\theta w = \sum_{n \in \bar{\Gamma}^+} \int_{\{N=n\}} w_n dP_\theta^{(n)} \quad , \quad \theta \in \Theta \quad ,$$

where $P_\theta^{(n)}$ denotes the restriction of P_θ from $(\Omega, \mathfrak{F})$ to $(\Omega, \mathfrak{F}_n)$, $n \in \bar{\Gamma}^+$.
One of the most important characteristics for assessing the properties of a test is the power function or its counterpart, the operating characteristic function.

D e f i n i t i o n 1.1.3. Let (N, δ) be any given test. Then the functions

$$M(\theta) = E_\theta \delta \chi_{\{N<\infty\}} \quad , \theta \in \Theta$$

and

$$Q(\theta) = E_\theta(1-\delta)\chi_{\{N<\infty\}} \quad , \theta \in \Theta$$

are called <u>power function</u> and <u>operating characteristic function</u> (OC-function) of (N,δ), respectively.

According to this definition, the power function provides for every $\theta \in \Theta$ the probability of acceptance of hypothesis H_1 by the test (N,δ) at a finite sampling stage. Evidently, we have

$$M(\theta) + Q(\theta) \leq 1 \quad , \theta \in \Theta ,$$

and (N,δ) is closed iff $M(\theta) + Q(\theta) = 1$ for every $\theta \in \Theta$.

If (N,δ) is a sequential test, the moments $E_\theta N^r$, $\theta \in \Theta$, $r \in \Gamma^+$, of the sample size N are important characteristics describing certain properties of the sample size. In this context it is usual to denote the first moment $E_\theta N$ as a function of θ as the <u>average sample number function</u> (ASN-function).

If (N,δ) is a test according to Definition 1.1.1, then sample size N is a so-called non-randomized sample size. Of course, it would be possible to admit also randomized sample sizes. One way to obtain such a sample size ist the following. As stated above, a sample size N w.r.t. $\{\mathcal{F}_n\}_{n\in\Gamma^+}$ can be represented in the form (1.1). This implies the following generalization. Let $\{\varphi_n\}_{n\in\Gamma^+}$ be a sequence of $(\mathcal{F}_n, \mathcal{B}^1)$-measurable functions $\varphi_n : \Omega \to [0,1]$, $n \in \Gamma^+$, and let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of Bernoulli distributed random variables, where

$$P_\theta(T_n=1 \mid \varphi_n=\varphi') = \varphi' \text{ and } P_\theta(T_n=0 \mid \varphi_n=\varphi') = 1 - \varphi'$$

for every $\theta \in \Theta$, $n \in \Gamma^+$ and $\varphi' \in [0,1]$. Then we get a randomized sample size N if N is defined by

$$N = \begin{cases} \inf\{n \geq 1 : T_n = 1\} & \text{, if such an n exists,} \\ \infty & \text{, otherwise,} \end{cases}$$

where the decision for stopping or continuing our experiment at the n^{th} sampling stage will depend on the result of an auxiliary experiment, including the observation of the random variable T_n. The properties of randomized sample sizes have been investigated in

detail by BAHADUR [8] , DÖHLER [26] and HECKENDORFF [40] . Due to the fact that no examples have been known in connection with tests so far, where such generalizations are of practical importance, we will renounce the consideration of such sample sizes here.

1.2 Tests based on a sequence of random variables

Frequently we have experiments where we can observe a sequence of random variables $\{X_n\}_{n\in\Gamma^+}$, defined on $(\Omega, \mathfrak{F})$, and where our decisions can only refer to these random variables. Then the sequence $\{X_n\}_{n\in\Gamma^+}$ generates a non-decreasing sequence $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ of sub-σ-algebras of $\mathfrak{F}$ with $\mathfrak{F}_n = \sigma(\vec{X}_n)$, $\vec{X}_n = (X_1,\dots,X_n)$, $n\in\Gamma^+$, so that $\mathfrak{F}_n$ contains the whole information of our experiment which is available, observing the vector $\vec{X}_n$, $n\in\Gamma^+$. If (N,δ) is a test based on $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ with $\mathfrak{F}_n = \sigma(X_n)$ for $n\in\Gamma^+$, we shall say (N,δ) <u>is a test based on</u> $\{X_n\}_{n\in\Gamma^+}$. For such tests the following assertion holds.

<u>L e m m a 1.2.1</u>. A test (N,δ) is based on $\{X_n\}_{n\in\Gamma^+}$ iff a sequence of Borel sets $\{B_n\}_{n\in\Gamma^+}$, $B_n\in\mathfrak{B}^n$, $n\in\Gamma^+$, exists so that

$$N = \begin{cases} \inf\{n\geq 1:\ \vec{X}_n\in B_n\}, & \text{if such an n exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{1.3}$$

holds.

P r o o f. (i) If $\{B_n\}_{n\in\Gamma^+}$ is any given sequence of Borel sets $B_n\in\mathfrak{B}^n$, $n\in\Gamma^+$, then, evidently (N,δ), if N is defined by (1.3), is a test based on $\{X_n\}_{n\in\Gamma^+}$.

(ii) Let (N,δ) be a test based on $\{X_n\}_{n\in\Gamma^+}$. Then, for every $n\in\Gamma^+$, we have $\{N = n\}\in\mathfrak{F}_n = \sigma(\vec{X}_n)$, and there exists a Borel set $B_n\in\mathfrak{B}^n$ with $\{N = n\} = \{\vec{X}_n\in B_n\}$, and $\vec{X}_n\in B_n$ implies $\chi_{\{N=n\}} = 1$. Therefore, we obtain

$$\inf\{n\geq 1:\ \vec{X}_n\in B_n\} = \inf\{n\geq 1:\ \chi_{\{N=n\}} = 1\}$$

so that (1.3) holds, and the proof is complete. ■

We remark that different sequences of Borel sets do not necessarily provide different sample sizes if the sample size is defined by (1.3). If namely $\{B_n\}_{n\in\Gamma^+}$ is any given sequence of Borel sets, $B_n\in\mathfrak{B}^n$, $n\in\Gamma^+$, and if N is defined by (1.3), then we have

$$\{N = n\} = \{\vec{X}_1 \notin B_1, \dots, \vec{X}_{n-1} \notin B_{n-1}, \vec{X}_n \in B_n\}$$
$$\subseteq \{\vec{X}_n \in B_n\}, \quad n \in \Gamma^+.$$

Otherwise, it is always possible to reduce a given set of Borel sets $\{B_n\}_{n\in\Gamma^+}$ to a set of Borel sets $\{B'_n\}_{n\in\Gamma^+}$ so that

$$\inf\{n \geq 1 : \vec{X}_n \in B_n\} = \inf\{n \geq 1 : \vec{X}_n \in B'_n\},$$

and

$$\{N = n\} = \{\vec{X}_n \in B'_n\}, \quad n \in \Gamma^+,$$

holds. In doing this, let B'_n be defined by

$$B'_n = B_n - \bigcup_{k=1}^{n-1} (B'_k \times (\bigtimes_{i=k+1}^{n} R_i^1)), \; R_i^1 = R^1 \text{ for } i \in \Gamma^+, \; n \in \Gamma^+, \tag{1.4}$$

then we have $N = n$ iff $\vec{X}_n \in B'_n$, $n \in \Gamma^+$. Based on the sequence $\{B'_n\}_{n\in\Gamma^+}$ defined by (1.4), we can obtain a partition $\{B'^{\infty}_n\}_{n\in\Gamma^+}$ of R^∞ in disjoint sets with finite bases given by

$$B'^{\infty}_n = B'_n \times (\bigtimes_{i=n+1}^{\infty} R_i^1), \; R_i^1 = R^1 \text{ for } i \in \Gamma^+, \; n \in \Gamma^+,$$

and

$$B'^{\infty}_\infty = R^\infty - \bigcup_{n\in\Gamma^+} B'^{\infty}_n .$$

Such a concept has been used by ARROW, BLACKWELL, GIRSHICK [7], describing sample sizes for sequential procedures.
An **important** special case is given by the following example.

<u>E x a m p l e 1.2.1</u>. The i.i.d. case. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of independent and identically distributed (i.i.d.) random variables with values in a measurable space $(\mathfrak{X}, \mathfrak{A})$. Denote by $P_\theta^{X_i}$ the corresponding distribution of X_i, $\theta \in \Theta$, $i \in \Gamma^+$. Then we can choose $\Omega, \mathfrak{F}, \{\mathfrak{F}_n\}_{n\in\Gamma^+}$ and $\mathfrak{P} = \{P_\theta, \theta \in \Theta\}$ in the following manner.

$$\Omega = \bigtimes_{i=1}^{\infty} \mathfrak{X}_i, \quad \mathfrak{X}_i = \mathfrak{X} \quad \text{for } i \in \Gamma^+,$$

$$\mathfrak{F} = \bigotimes_{i=1}^{\infty} \mathfrak{A}_i, \quad \mathfrak{A}_i = \mathfrak{A} \quad \text{for } i \in \Gamma^+,$$

$$\mathfrak{F}_n = \bigotimes_{i=1}^{n} \mathfrak{A}_i \times (\bigtimes_{j=n+1}^{\infty} \mathfrak{X}_j), \quad \mathfrak{A}_i = \mathfrak{A} \text{ for } i \in \Gamma^+,$$
$$\mathfrak{X}_j = \mathfrak{X} \text{ for } j \in \Gamma^+,$$

$$P_\theta = \bigotimes_{i=1}^{\infty} P_\theta^{X_i}, \quad P_\theta^{X_i} = P_\theta^{X_1} \quad \text{for } i \in \Gamma^+,$$

where for every $i \in \Gamma^+$ the random variable X_i can be assumed as a random variable $X_i : \Omega \longrightarrow \mathfrak{X}$ with

$$X_i(\omega) = x_i \quad \text{for} \quad \omega = (x_1,\ldots,x_i,\ldots) \in \Omega .$$

If the $\{X_n\}_{n\in\Gamma^+}$ are random variables having a density $f_\theta(x)$ w.r.t. some measure μ on $(\mathfrak{X}, \mathfrak{A})$, then we have $P_\theta^{(n)} \ll \mu^{*(n)}$, where $P_\theta^{(n)}$ and $\mu^{*(n)}$ denote the restrictions of P_θ and $\mu^* = \bigotimes_{i=1}^{\infty} \mu_i$, $\mu_i = \mu$ for $i \in \Gamma^+$, from $(\Omega, \mathfrak{F})$ to $(\Omega, \mathfrak{F}_n)$, $n \in \Gamma^+$. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables, we obtain for the Radon-Nikodym-derivative of $P_\theta^{(n)}$ w.r.t. $\mu^{*(n)}$

$$\frac{dP_\theta^{(n)}}{d\mu^{*(n)}}(\omega) = \prod_{i=1}^{n} f_\theta(X_i(\omega)) , \quad \omega = (x_1,\ldots,x_n,\ldots) \in \Omega .$$

It is usual to denote this derivative as the <u>likelihood function</u> at sample size n. ■

1.3 Tests based on a sequence of statistics

However, the structure of a test (N, δ) based on $\{X_n\}_{n\in\Gamma^+}$ may be rather complicated. This essentially depends on the geometrical properties of the corresponding sequence of Borel sets $\{B_n\}_{n\in\Gamma^+}$, defining sample size N. For instance, it can be very cumbersome to decide whether a given sampling vector X_n belongs to B_n or does not. Therefore, in view of the practical implementation of a test (N, δ), further simplification of its structure is necessary. This can be done if we consider, instead of the sequence of random variables $\{X_n\}_{n\in\Gamma^+}$, a sequence of statistics. If we have a sequence of vector valued statistics then we will assume that for every $n \in \Gamma^+$ the dimension is equal.

Let $(\Omega, \mathfrak{F}, \mathfrak{P})$ be a given statistical structure, let $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ be a sequence of non-decreasing sub-σ-algebras of $\mathfrak{F}$, and let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of statistics $T_n : \Omega \longrightarrow R^k$, $k \in \Gamma^+$, with $\mathfrak{G}_n = \sigma(T_n) \subseteq \mathfrak{F}_n$, $n \in \Gamma^+$. Based on the sequence $\{T_n\}_{n\in\Gamma^+}$, we can obtain a sample size as follows. Let $\{C_n\}_{n\in\Gamma^+}$ be any given sequence of Borel sets $C_n \in \mathfrak{B}^k$, $n \in \Gamma^+$, then

$$N = \begin{cases} \inf \{n \geq 1 : T_n \in C_n\}, & \text{if such an n exists,} \\ \infty & \text{, otherwise} \end{cases} \tag{1.5}$$

is a sample size w.r.t. $\{\mathcal{F}_n\}_{n\in\Gamma^+}$. Let now $\{d_n\}_{n\in\Gamma^+}$ be a sequence of measurable functions $d_n: R^k \to [0,1]$, $n \in \Gamma^+$, then $d_n(T_n)$ is $(\mathcal{G}_n, \mathcal{B}^1)$-measurable. Since $\mathcal{G}_n = \sigma(T_n) \subseteq \mathcal{F}_n$ for every $n \in \Gamma^+$, the function

$$\delta = \sum_{n\in\bar{\Gamma}^+} d_n(T_n)\, \chi_{\{N=n\}} \tag{1.6}$$

is $(\mathcal{F}_N, \mathcal{B}^1)$-measurable, and δ is a terminal decision rule in the sense of Definition 1.1.1. A test (N, δ) where N and δ are defined according to (1.5) and (1.6) is called a <u>test based on the sequence of statistics</u> $\{T_n\}_{n\in\Gamma^+}$. We shall see that the tests considered in Sections 2 and 3 are just of such a simple structure.

Beside the simplification of the structure of the sample size and the terminal decision rule, sequences of statistics will also play an important role in connection with the computation of the characteristics of any given test. In this context, the question arises whether sequences of statistics exist where it is sufficient to refer only to these sequences in computing the characteristics.

To answer this question let $\{T_n\}_{n\in\Gamma^+}$ be any given sequence of statistics and $\{\mathcal{G}_n\}_{n\in\Gamma^+}$ the corresponding sequence of σ-algebras with $\mathcal{G}_n = \sigma(T_n) \subseteq \mathcal{F}_n$, $n \in \Gamma^+$. Then the σ-algebra $\mathcal{G}_n$ contains only those events of our experiment which can be observed by means of an observation of the statistic T_n, $n \in \Gamma^+$. We note that then, in general, the sequence $\{\mathcal{G}_n\}_{n\in\Gamma^+}$ will not be increasing monotonously. Let $\mathcal{G}_N$ be a σ-algebra analogously to $\mathcal{F}_N$ defined by

$$\mathcal{G}_N = \sigma\{\{N = n\} \cap G_n : G_n \in \mathcal{G}_n,\ n \in \bar{\Gamma}^+\},\quad \mathcal{G}_\infty = \{\emptyset, \Omega\}.$$

Since $\mathcal{G}_n \subseteq \mathcal{F}_n$ for $n \in \bar{\Gamma}^+$, we have $\mathcal{G}_N \subseteq \mathcal{F}_N$ so that $\mathcal{G}_N$ is, in general, only a sub-σ-algebra of the σ-algebra $\mathcal{F}_N$ of the N-past. Furthermore, one can verify that

$$\mathcal{G}_N = \sigma(N, T_N) \text{ with } T_N = \sum_{n\in\bar{\Gamma}^+} T_n\, \chi_{\{N=n\}},\quad T_\infty = 0$$

holds (see e.g. [26], p. 24).
Let now $w: \Omega \to R^1$ be an $(\mathcal{F}_N, \mathcal{B}^1)$-measurable and P_Θ-integrable

function, $\theta \in \Theta$, then we obtain by definition of the conditional expectation

$$E_\theta w = E_\theta(E_\theta(w \mid \mathcal{G}_N)).$$

Since $\mathcal{G}_N = \sigma(N, T_N)$, a measurable function $\hat{w}_\theta : R^{k+1} \longrightarrow R^1$ exists with

$$E_\theta(w \mid \mathcal{G}_N) = \hat{w}_\theta(N, T_N), \qquad P_\theta \text{ - a.s.}$$

so that

$$E_\theta w = E_\theta \hat{w}_\theta(N, T_N)$$

holds. This implies any characteristic $E_\theta w$ of a test (N, δ) w.r.t. $\{\mathcal{F}_n\}_{n \in \Gamma^+}$ can be represented as an expectation value of a certain function $\hat{w}_\theta$, depending, beside the parameter θ, only on N and T_N.

It is somewhat inconvenient in this representation that, as a rule, the function w_θ depends on parameter $\theta \in \Theta$. If, for instance, (N, δ) is a test based on $\{X_n\}_{n \in \Gamma^+}$ this dependence can be interpreted as follows. It compensates the arising loss of information w.r.t. parameter θ if we refer only to sequence $\{T_n\}_{n \in \Gamma^+}$ instead of $\{X_n\}_{n \in \Gamma^+}$. Therefore, in connection with the computation of a characteristic of a test (N, δ), such sequences of statistics $\{T_n\}_{n \in \Gamma^+}$ are of special interest where a version $E(w \mid \mathcal{G}_N)$ of the conditional expectation $E_\theta(w \mid \mathcal{G}_N)$ exists which does not depend on θ for $\theta \in \Theta$.

<u>D e f i n i t i o n 1.3.1</u>. The σ-algebra $\mathcal{G}_N = \sigma(N, T_N)$ or the statistic (N, T_N), respectively, is said to be <u>sufficient for</u> $\mathcal{F}_N$ and Θ, if for every $(\mathcal{F}_N, \mathcal{B}^1)$-measurable and P_θ-integrable function $w : \Omega \longrightarrow R^1$, a version $E(w \mid \mathcal{G}_N)$ exists, which does not depend on $\theta \in \Theta$.

Evidently, $\mathcal{G}_N$ is sufficient for $\mathcal{F}_N$ and Θ iff

$$\int_G E_\theta(w \mid \mathcal{G}_N) dP_\theta = \int_G E(w \mid \mathcal{G}_N) dP_\theta, \quad G \in \mathcal{G}_N, \quad \theta \in \Theta \tag{1.7}$$

holds. We remark that this relation is satisfied if a version $P(A \mid \mathcal{G}_N)$ of the conditional probability $P_\theta(A \mid \mathcal{G}_N)$, $A \in \mathcal{F}$, exists so that

$$P(A \mid \mathcal{G}_N) = P_\theta(A \mid \mathcal{G}_N), \qquad P_\theta \text{ - a.s.}, \ \theta \in \Theta,$$

holds. That means the requirement for sufficiency is ultimately a condition on the conditional probability $P_\theta(A \mid \mathcal{G}_N)$. An immediate consequence of the sufficiency is the following lemma now:

L e m m a 1.3.1. Let (N, δ) be a closed test based on $\{\mathcal{F}_n\}_{n\in\Gamma^+}$. Let $w: \Omega \longrightarrow R^1$ be an $(\mathcal{F}_N, \mathcal{B}^1)$-measurable and P_θ-integrable function, $\theta \in \Theta$, and let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of statistics where $\mathcal{G}_N = \sigma(N, T_N)$ is sufficient for $\mathcal{F}_N$ and Θ. Then a $(\mathcal{G}_N, \mathcal{B}^1)$-measurable function $\hat{w}: \Omega \longrightarrow R^1$ and a sequence $\{v_n\}_{n\in\Gamma^+}$ of measurable functions $v_n: R^k \longrightarrow R^1$, $n \in \Gamma^+$, exists with

$$E_\theta w = E_\theta \hat{w} \quad , \theta \in \Theta \tag{1.8}$$

and

$$\hat{w} = \sum_{n\in\Gamma^+} v_n(T_n)\chi_{\{N=n\}}, \ P_\theta - \text{a.s.}, \ \theta \in \Theta . \tag{1.9}$$

P r o o f: Since $\mathcal{G}_N$ is sufficient for $\mathcal{F}_N$ and Θ, a version $E(w|\mathcal{G}_N)$ of the conditional expectation $E_\theta(w|\mathcal{G}_N)$ exists which does not depend on θ for $\theta \in \Theta$. With $\hat{w} = E(w|\mathcal{G}_N)$ we obtain (1.8). Now, $\hat{w}$ as a version of $E_\theta(w|\mathcal{G}_N)$ is a $(\mathcal{G}_N, \mathcal{B}^1)$-measurable function. Then a sequence $\{\hat{w}_n\}_{n\in\bar{\Gamma}^+}$ of $(\mathcal{G}_n, \mathcal{B}^1)$-measurable functions $\hat{w}_n: \Omega \longrightarrow R^1$, $n \in \Gamma^+$, exists with

$$\hat{w} = \sum_{n\in\bar{\Gamma}^+} \hat{w}_n \chi_{\{N=n\}} .$$

Since (N, δ) is closed, we obtain

$$\hat{w} = \sum_{n\in\Gamma^+} \hat{w}_n \chi_{\{N=n\}} , \ P_\theta - \text{a.s.}, \ \theta \in \Theta . \tag{1.10}$$

Finally, for every $n \in \Gamma^+$ the σ-algebra $\mathcal{G}_n$ is generated by T_n. Then for every $n \in \Gamma^+$ a measurable function $v_n: R^k \to R^1$ exists with

$$\hat{w}_n = v_n(T_n) , \ P_\theta - \text{a.s.}, \ \theta \in \Theta .$$

This, together with (1.10), implies (1.9). ■

C o r o l l a r y 1.3.1. Let (N, δ) be a closed test based on $\{\mathcal{F}_n\}_{n\in\Gamma^+}$, let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of statistics so that $\mathcal{G}_N = \sigma(N, T_N)$ is sufficient for $\mathcal{F}_N$ and Θ. Then a $(\mathcal{G}_N, \mathcal{B}^1)$-measurable terminal decision rule $\hat{\delta}$ and a sequence of measurable functions $\{d_n\}_{n\in\Gamma^+}$, $d_n: R^k \longrightarrow [0,1]$, $n \in \Gamma^+$, exist with

$$M(\theta) = E_\theta \delta = E_\theta \hat{\delta} \quad , \theta \in \Theta , \tag{1.11}$$

and

$$\hat{\delta} = \sum_{n\in\Gamma^+} d_n(T_n)\chi_{\{N=n\}}, \ P_\theta - \text{a.s.}, \ \theta \in \Theta . \tag{1.12}$$

P r o o f. By Lemma 1.3.1 a $(\mathfrak{G}_N, \mathfrak{L}^1)$-measurable terminal decision rule $\hat{\delta}$ and a corresponding sequence $\{d_n\}_{n \in \Gamma^+}$ exist so that $E_\theta \delta = E_\theta \hat{\delta}$ for $\theta \in \Theta$, and (1.12) holds. Since (N, δ) is closed, we further have

$$M(\theta) = E_\theta \delta \chi_{\{N < \infty\}} = E_\theta \delta \text{ and } E_\theta \hat{\delta} \chi_{\{N < \infty\}} = E_\theta \hat{\delta} ,$$

$\theta \in \Theta$. That implies (1.11).■

The corollary shows that for a closed test in case of the sufficiency of $\mathfrak{G}_N$ we can restrict our attention to those terminal decision rules which depend on the set $\{N = n\}$ only on $T_n(\omega)$, $n \in \Gamma^+$. Indeed, the sufficiency of $\mathfrak{G}_N$ does not imply, in general, that also the sample size N can be represented only by means of the sequence $\{T_n\}_{n \in \Gamma^+}$.

A further consequence of Lemma 1.3.1 in regard of a practical computation of the characteristics of a test is the following.

C o r o l l a r y 1.3.2. Suppose the assumptions of Lemma 1.3.1 are fulfilled. Then there exists a measurable function $\hat{w}^*$ which does not depend on θ for $\theta \in \Theta$, where

$$E_\theta w = E_\theta \hat{w}^*(N, T_N) , \quad \theta \in \Theta . \tag{1.13}$$

P r o o f. The function $\hat{w}$ introduced by Lemma 1.3.1 is $(\mathfrak{G}_N, \mathfrak{L}^1)$-measurable. Because of $\mathfrak{G}_N = \mathfrak{S}(N, T_N)$ a measurable function $\hat{w}^*$ exists with

$$\hat{w} = \hat{w}^*(N, T_N) , \quad P_\theta \text{ - a.s.} , \quad \theta \in \Theta .$$

This, together with (1.8), implies (1.13).■

Hence, if the assumptions of Lemma 1.3.1 are fulfilled, the characteristics of any given test (N, δ) can be represented as expectation values of certain functions only depending on N and T_N. That means that in case of the sufficiency of $\mathfrak{G}_N$ the investigation into the statistical properties of a test (N, δ) by means of its characteristics can be reduced to the consideration of expectation values of functions which only depend on N and T_N so that methods of computing such expectation values will play a special role.

The computation of an expectation value $E_\theta w(N, T_N)$ would be possible, for instance, by means of the common distribution function of (N, T_N). For some special cases we can obtain representation formulas for the distribution function of (N, T_N). Such representation formulas have

been investigated by DE GROOT [24], TRYBULA [75] and FRANZ, WINKLER [33], but even in these cases an explicite determination of the distribution function of (N,T_N) is very cumbersome. By means of representation formulas FRANZ, WINKLER [33] obtained for certain functions $\hat{w}^*$ relations between $E_\theta \hat{w}^*(N,T_N)$ and the moments of N and T_N, so that the computation of some characteristic $E_\theta \hat{w}^*(N,T_N)$ may be reduced to the computation of certain moments of N and T_N. Special cases of such moment equations, considered in [33], are WALD's equation and the equations of HALL [39].

In Section 3, we shall return to the computation of the characteristics of a test (N,δ). There a quite general method is developed, allowing us the characteristics of WALD's likelihood ratio test to be computed, based on a sequence of discrete random variables without any explicite reference to the distribution function of (N,T_N).

1.4 Sufficient sequences of statistics

It follows from Section 1.3 that the sufficiency of $\mathfrak{G}_N$ or (N,T_N), respectively, for $\mathfrak{F}_N$ and Θ plays an essential role in regard of a simplification of the structure of a test (N,δ) as well as the computation of its characteristics. Therefore, here we consider some criteria for the sufficiency adapted to sequential tests.

Let N be a sample size w.r.t. $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$, and let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of statistics so that $\mathfrak{G}_n = \sigma(T_n) \subseteq \mathfrak{F}_n$, $n\in\Gamma^+$. Then, by HECKENDORFF [40], Theorem 1.10, the following assertion holds. If N is closed and $\mathfrak{G}_n$ is sufficient for $\mathfrak{F}_n$ and Θ for every $n\in\Gamma^+$, then (N,T_N) with

$$T_N = \sum_{n\in\Gamma^+} T_n \chi_{\{N=n\}}$$

is sufficient for $\mathfrak{F}_N$ and Θ. Thus, the investigation into the sufficiency of (N,T_N) can be reduced to the investigation into the sufficiency of T_n for every $n\in\Gamma^+$. In particular, the following lemma holds.

Lemma 1.4.1. Let N be a closed sample size w.r.t. $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$, $\Theta = \{\theta_0,\theta_1\}$, $\theta_0 \neq \theta_1$. Suppose that $P_{\theta_1}^{(n)} \ll P_{\theta_0}^{(n)}$ for every $n\in\Gamma^+$. Denote by L_{n,θ_0,θ_1} any version of the Radon-Nikodym-derivative of $P_{\theta_1}^{(n)}$ w.r.t $P_{\theta_0}^{(n)}$ $n\in\Gamma^+$. Then $(N,L_{N,\theta_0,\theta_1})$ with

$$L_{N,\theta_o,\theta_1} = \sum_{n\in\Gamma^+} L_{n,\theta_o,\theta_1}\chi_{\{N=n\}}$$

is sufficient for $\mathcal{F}_N$ and Θ.

P r o o f. Let $\mathcal{G}_n$ be defined by $\mathcal{G}_n = \sigma(L_{n,\theta_o,\theta_1})$, $n\in\Gamma^+$. Denote by $E_\theta w$ the expectation value of a certain function w w.r.t. the measure $P_\theta^{(n)}$. Then, for any given $F\in\mathcal{F}_n$ and $G\in\mathcal{G}_n$ we have

$$\begin{aligned}
P_{\theta_1}^{(n)}(F\cap G) &= E_{\theta_1}\chi_{F\cap G}\\
&= E_{\theta_o}(\chi_{F\cap G}L_{n,\theta_o,\theta_1})\\
&= E_{\theta_o}(E_{\theta_o}(\chi_{F\cap G}L_{n,\theta_o,\theta_1}|\mathcal{G}_n))\\
&= E_{\theta_o}(\chi_G L_{n,\theta_o,\theta_1} E_{\theta_o}(\chi_F|\mathcal{G}_n)).
\end{aligned}$$

Otherwise, we obtain

$$\begin{aligned}
P_{\theta_1}^{(n)}(F\cap G) &= E_{\theta_1}(E_{\theta_1}(\chi_{F\cap G}|\mathcal{G}_n))\\
&= E_{\theta_1}(\chi_G E_{\theta_1}(\chi_F|\mathcal{G}_n))\\
&= E_{\theta_o}(\chi_G L_{n,\theta_o,\theta_1} E_{\theta_1}(\chi_F|\mathcal{G}_n)).
\end{aligned}$$

This implies for every $F\in\mathcal{F}_n$

$$E_{\theta_o}(\chi_F|\mathcal{G}_n) = E_{\theta_1}(\chi_F|\mathcal{G}_n)\ ,\ P_{\theta_o}^{(n)}\text{ - a.s.}$$

and

$$P_{\theta_o}^{(n)}(F|\mathcal{G}_n) = P_{\theta_1}^{(n)}(F|\mathcal{G}_n)\ ,\ P_{\theta_o}^{(n)}\text{ - a.s.}$$

respectively. That means, L_{n,θ_o,θ_1} is sufficient for $\mathcal{F}_n$ and $\Theta = \{\theta_o,\theta_1\}$ for every $n\in\Gamma^+$. N is closed, applying Theorem 1.10 of [40], we obtain $(N, L_{N,\theta_o,\theta_1})$ is sufficient for $\mathcal{F}_N$ and $\Theta = \{\theta_o,\theta_1\}$.

It is common practice to denote any version of the Radon-Nikodym-derivative $dP_{\theta_1}^{(n)}/dP_{\theta_o}^{(n)}$ of $P_{\theta_1}^{(n)}$ w.r.t $P_{\theta_o}^{(n)}$ as a <u>likelihood ratio at the sample size</u> n, $n\in\Gamma^+$. This notation is also motivated by the following example.

E x a m p l e 1.4.1. Let $\{X_n\}_{n \in \Gamma^+}$ be a sequence of i.i.d. random variables with a density $f_\theta(x)$.[1] Referring to Example 1.2.1, we obtain $P_{\theta_0}^{(n)} \ll \mu^{*(n)}$ and $P_{\theta_1}^{(n)} \ll \mu^{*(n)}$ for $n \in \Gamma^+$. Since the random variables $\{X_n\}_{n \in \Gamma^+}$ are assumed to be independent, we obtain for the corresponding densities

$$\frac{dP_\theta^{(n)}}{d\mu^{*(n)}}(\omega) = \prod_{i=1}^{n} f_\theta(X_i(\omega)), \ \theta \in \{\theta_0, \theta_1\}, \omega \in \Omega . \tag{1.14}$$

If additionally $P_{\theta_1}^{(n)} \ll P_{\theta_0}^{(n)}$, $n \in \Gamma^+$, holds, then $P_{\theta_1}^{(n)} \ll P_{\theta_0}^{(n)} \ll \mu^{*(n)}$ implies

$$\frac{dP_{\theta_1}^{(n)}}{d\mu^{*(n)}} = \frac{dP_{\theta_1}^{(n)}}{dP_{\theta_0}^{(n)}} \frac{dP_{\theta_0}^{(n)}}{d\mu^{*(n)}} = L_{n,\theta_0,\theta_1} \cdot \frac{dP_{\theta_0}^{(n)}}{d\mu^{*(n)}} \ , \ \mu^{*(n)}\text{- a.s.}$$

This and (1.14) implies that we obtain a version L_{n,θ_0,θ_1} of the Radon-Nikodym-derivative $dP_{\theta_1}^{(n)}/dP_{\theta_0}^{(n)}$ if we put

$$L_{n,\theta_0,\theta_1} = \prod_{i=1}^{n} \frac{f_{\theta_1}(X_i)}{f_{\theta_0}(X_i)}$$

where in case of $\prod_{i=1}^{n} f_{\theta_0}(X_i(\omega)) = 0$ we can choose an arbitrary value for $L_{n,\theta_0,\theta_1}(\omega)$.■

If now (N, δ) is any given test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ and if $(N, L_{N,\theta_0,\theta_1})$ is sufficient for $\mathfrak{F}_N$ and Θ in sense of Lemma 1.4.1, then, according to Corollary 1.3.1, we can restrict our attention to those terminal decision rules $\hat{\delta}$ which depend on $\{N=n\}$ only on L_{n,θ_0,θ_1}, $n \in \Gamma^+$. As already stated above in connection with Corollary 1.3.1, the sufficiency of $(N, L_{N,\theta_0,\theta_1})$ does not imply, as a rule, that also sample size N can be represented only by means of the sequence $\{L_{n,\theta_0,\theta_1}\}_{n \in \Gamma^+}$. But we shall see in Section 2 that tests whose sample size can be represented as

$$N = \begin{cases} \inf \{n \geq 1: L_{n,\theta_0,\theta} \in C_n\} & \text{, if such an n exists,} \\ \infty & \text{, otherwise,} \end{cases}$$

[1] This terminology is used to express that $\{X_n\}_{n \in \Gamma^+}$ is a sequence of i.i.d. random variables having a density $f_\theta(x)$ w.r.t some measure μ on $(\mathfrak{X}, \mathfrak{A})$.

where $\{C_n\}_{n\in\Gamma^+}$ denotes a given sequence of Borel sets $C_n \in \mathcal{B}^1$, $n\in\Gamma^+$, will have quite far-reaching optimality properties.

The following lemma now presents a modification of NEYMAN's well-known factorization criterion for sufficiency adapted to sequential tests.

<u>L e m m a 1.4.2</u>. Let N be a closed sample size w.r.t. $\{\mathcal{F}_n\}_{n\in\Gamma^+}$, let $\{T_n\}_{n\in\Gamma^+}$ be a sequence of statistics with $\mathcal{T}_n = \sigma(T_n) \subseteq \mathcal{F}_n$, $n\in\Gamma^+$. If for every pair $\theta', \theta'' \in \Theta$, $\theta \neq \theta''$ and $n\in\Gamma^+$ a measurable function $g_{n,\theta',\theta''}$ exists so that

$$L_{n,\theta',\theta''} = g_{n,\theta',\theta''}(T_n) \ , \quad P_\theta^{(n)} \text{ - a.s. ,} \tag{1.15}$$

then (N,T_N) with $T_N = \sum_{n\in\Gamma^+} T_n \chi_{\{N=n\}}$ is sufficient for $\mathcal{F}_N$ and Θ.

P r o o f. For any given pair $\theta',\theta'' \in \Theta$, $\theta' \neq \theta''$ and $n\in\Gamma^+$ let $\mathcal{G}_n$ be defined by $\mathcal{G}_n = \sigma(L_{n,\theta',\theta''})$. By (1.15) we obtain $\mathcal{G}_n \subseteq \mathcal{T}_n = \sigma(T_n)$ so that $L_{n,\theta',\theta''}$ is also $(\mathcal{T}_n, \mathcal{B}^1)$-measurable. Then, analogously to the proof of Lemma 1.4.1, we can show that for every $F \in \mathcal{F}_n$

$$P_{\theta'}^{(n)}(F\,|\,\mathcal{T}_n) = P_{\theta''}^{(n)}(F\,|\,\mathcal{T}_n) \ , \quad P_{\theta'}^{(n)} \text{ - a.s.,}$$

holds. Therefore, $\mathcal{T}_n$ or the statistic T_n is for every $n\in\Gamma^+$ sufficient for $\mathcal{F}_n$ and Θ. N is closed, applying Theorem 1.10 of [40], this completes the proof. ■

If Lemma 1.4.2 holds, then it is also usual to say that $\{T_n\}_{n\in\Gamma^+}$ forms a sequence of sufficient statistics for $\{\mathcal{F}_n\}_{n\in\Gamma^+}$ and Θ. Important examples of distribution families where sequences of sufficient statistics exist are the exponential families.

<u>E x a m p l e 1.4.2</u>. Let $\{P_\theta,\ \theta\in\Theta\}$ be a family of probability measures on $(\Omega,\mathcal{F})$, and let $\{\mathcal{F}_n\}_{n\in\Gamma^+}$ be a sequence of non-decreasing sub-σ-algebras of $\mathcal{F}$. A family of non-degenerated probability measures $\{P_\theta^{(n)},\ \theta\in\Theta\}$ on $(\Omega,\mathcal{F}_n)$, $n\in\Gamma^+$, dominated by a σ-finite measure μ on $(\Omega,\mathcal{F}_n)$, is said to be an <u>exponential family</u> if real-valued functions $c_n:\Theta \longrightarrow R^1$, $c_n > 0$, and $d_n^{(j)}:\Theta \longrightarrow R^1$ and $(\mathcal{F}_n,\mathcal{B}^1)$-measurable functions $h_n:\Omega \longrightarrow R^1$ and $T_n^{(j)}:\Omega \longrightarrow R^1$ $j = 1,\dots,k$, $k\in\Gamma^+$ exist so that for every $\theta\in\Theta$

$$\frac{dP_\theta^{(n)}}{d\mu^{(n)}}(\omega) = c_n(\theta)\exp\Big(\sum_{j=1}^{k} d_n^{(j)}(\theta)T_n^{(j)}(\omega)\Big)\, h_n(\omega)\ , \quad \mu^{(n)} \text{ - a.s.}$$

If $\{P_\theta^{(n)}, \theta \in \Theta\}$ is an exponential family, then any two measures $P_{\theta'}^{(n)}$ and $P_{\theta''}^{(n)}$ of this family are equivalent measures, and we obtain for the likelihood ratio

$$L_{n,\theta',\theta''} = \frac{c_n(\theta'')}{c_n(\theta')} \exp \left(\sum_{j=1}^{k} (d_n^{(j)}(\theta'') - d_n^{(j)}(\theta')) \cdot T_n^{(j)}\right) ,$$

$\mu^{(n)}$ - a.s. This implies, the likelihood ratio has the form (1.15) with $T_n = (T_n^{(1)},\dots,T_n^{(k)})$ and for every $n \in \Gamma^+$ the statistic T_n is sufficient for $\mathcal{F}_n$ and Θ. Hence, (N,T_N) is for every closed N sufficient for $\mathcal{F}_N$ and Θ. An important special case is again the i.i.d.-case of Example 1.2.1. We suppose that the family $\mathcal{P}^X = \{P_\theta^X, \theta \in \Theta\}$ is an exponential family. Then real-valued functions $c:\Theta \to R^1$, $d_j:\Theta \to R^1$, $h:\mathfrak{X} \to R^1$ and $t_j:\mathfrak{X} \to R^1$, $j = 1,\dots,k$, exist so that density $f_\theta(x)$ of P_θ^X w.r.t. μ can be represented as

$$f_\theta(x) = c(\theta) \exp\left(\sum_{j=1}^{k} d_j(\theta)\ _j(\theta)\right) h(x) \ , \ x \in \mathfrak{X} \ , \ \theta \in \Theta \ .$$

Then we obtain for the likelihood ratio

$$L_{n,\theta',\theta''} = \left(\frac{c(\theta'')}{c(\theta')}\right)^n \exp\left(\sum_{j=1}^{k} ((d_j(\theta'')-d_j(\theta'))\sum_{i=1}^{n} t_j(X_i)\right), \ n \in \Gamma^+ .$$

Hence, $T_n = (T_1^{(n)},\dots,T_k^{(n)})$ with $T_j^{(n)} = \sum_{i=1}^{n} t_j(X_i)$, $j = 1,\dots,k$, is for every $n \in \Gamma^+$ sufficient for $\mathcal{F}_n$ and Θ , and again by Lemma 1.4.2, (N,T_N) is for every closed N sufficient for $\mathcal{F}_N$ and Θ .■

The sequence of statistics $\{T_n\}_{n \in \Gamma^+}$ considered in this example possesses for the i.i.d.-case the additional property that for every $n \in \Gamma^+$

$$T_j^{(n+1)} = T_j^{(n)} + t_j(X_{n+1}) \ , \ j = 1,\dots,k \ ,$$

holds. This is a special version of the so-called <u>transitivity</u> (see [40] , Section 1.5) and allows even more far-reaching data reduction than sufficiency alone. Then also sample size N can be represented only by means of the sequence $\{T_n\}_{n \in \Gamma^+}$ (see [40], Theorem 1.14). We notice, that the transitivity of the sequence $\{T_n\}_{n \in \Gamma^+}$ will be moreover an advantageous property in connection with the computation of the characteristics of a test.

1.5 Convergence properties of the likelihood ratio sequence

We have seen in the previous section that the sequence $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ of the likelihood ratios at the sample size n, $n\in\Gamma^+$, forms a sequence of statistics allowing a simplification of the structure of a test (N,δ) for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ in the sense of the sufficiency concept. Obtaining additional clues to the choice of the sample size and the terminal decision rule of a test (N,δ) for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ based on the sequence $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ we consider some convergence properties of this sequence for $n\to\infty$.

Lemma 1.5.1. Let $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ be the sequence of the likelihood ratios at sample size n, $n\in\Gamma^+$, then we have

$$P_{\theta_0}(\lim_{n\to\infty} L_{n,\theta_0,\theta_1} = 0) = 1 \text{ iff } \lim_{n\to\infty} E_{\theta_0} L^{\frac{1}{2}}_{n,\theta_0,\theta_1} = 0 \quad (1.16)$$

and

$$P_{\theta_1}(\lim_{n\to\infty} L^{-1}_{n,\theta_0,\theta_1} = 0) = 1 \text{ iff } \lim_{n\to\infty} E_{\theta_1} L^{-\frac{1}{2}}_{n,\theta_0,\theta_1} = 0 \quad (1.17),$$

respectively.

Proof. First we notice that the sequence $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ forms a martingal in respect of $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$ and P_{θ_0}. This follows from

$$E_{\theta_0}|L_{n,\theta_0,\theta_1}| = E_{\theta_0}L_{n,\theta_0,\theta_1} = 1 \quad \text{for} \quad n\in\Gamma^+$$

and

$$\int_A L_{n,\theta_0,\theta_1} dP_{\theta_0} = \int_A \frac{dP^{(n)}_{\theta_1}}{dP^{(n)}_{\theta_0}} dP^{(n)}_{\theta_0} = P^{(n)}_{\theta_1}(A)$$

$$= P^{(n+1)}_{\theta_1}(A) = \int_A L_{n+1,\theta_0,\theta_1} dP_{\theta_0} \quad \text{for } A\in\mathfrak{F}_n,\ n\in\Gamma^+$$

This martingal property and $L_{n,\theta_0,\theta_1} \geq 0$ for $n\in\Gamma^+$ implies

$$\lim_{n\to\infty} L_{n,\theta_0,\theta_1} = L_\infty \geq 0\ ,\ P_{\theta_0} - \text{a.s.},$$

where L_∞ denotes a P_{θ_0} - integrable random variable (see e.g. BAUER [11], § 60, Corollary 1). Further, because of $L_{n,\theta_0,\theta_1} \geq 0$ and $E_{\theta_0}L_{n,\theta_0,\theta_1} = 1$ for $n\in\Gamma^+$ we obtain for every $a > 0$

$$\int_{\{L^{\frac{1}{2}}_{n,\theta_0,\theta_1} > a\}} L^{\frac{1}{2}}_{n,\theta_0,\theta_1} \, dP_{\theta_0} \leqslant \int_{\Omega} \frac{L^{\frac{1}{2}}_{n,\theta_0,\theta_1}}{a} L^{\frac{1}{2}}_{n,\theta_0,\theta_1} \, dP_{\theta_0} = \frac{E_{\theta_0} L_{n,\theta_0,\theta_1}}{a}, \quad n \in \Gamma^+,$$

so that

$$\lim_{a\to\infty} \sup_n \int_{\{L^{\frac{1}{2}}_{n,\theta_0,\theta_1} > a\}} L^{\frac{1}{2}}_{n,\theta_0,\theta_1} \, dP_{\theta_0} = 0.$$

Hence, the sequence $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ is uniformly integrable w.r.t. P_{θ_0} and by Theorem 1.3 of CHOW et al. [22] we obtain

$$\lim_{n\to\infty} E_{\theta_0} L^{\frac{1}{2}}_{n,\theta_0,\theta_1} = E_{\theta_0} L^{\frac{1}{2}}_{\infty} . \tag{1.18}$$

Because of $L_\infty \geqslant 0$ we obtain

$$E_{\theta_0} L^{\frac{1}{2}}_{\infty} = 0 \quad \text{iff} \quad P_{\theta_0}(L_\infty = 0) = 1 .$$

This, together with (1.18), provides (1.16). In the same way we obtain (1.17).■

This lemma contains certain clues concerning the structure of a test (N,δ) for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on the sequence of likelihood ratios $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$. If this sequence satisfies

$$P_{\theta_0}(\lim_{n\to\infty} L_{n,\theta_0,\theta_1} = 0) = 1$$

and

$$P_{\theta_1}(\lim_{n\to\infty} L^{-1}_{n,\theta_0,\theta_1} = 0) = 1,$$

where the last condition is also equivalent to

$$P_{\theta_1}(\lim_{n\to\infty} L_{n,\theta_0,\theta_1} = \infty) = 1,$$

then it is obvious to require that the test (N,δ) possesses a structure which can be characterized as follows. We continue sampling for $n = 1,2,\ldots$ until we observe a sufficiently small or large value of the likelihood ratio L_{n,θ_0,θ_1}. After stopping sampling we accept the hypothesis H_0 if L_{n,θ_0,θ_1} is small and reject H_0 or accept H_1 if L_{n,θ_0,θ_1} is large, respectively. This is the basic

structure of the so-called likelihood ratio tests considered in details in Section 2.
The subsequent lemma now presents a further criterion for (1.16) and (1.17), respectively, for the i.i.d.-case.

<u>L e m m a 1.5.2</u>. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having a density $f_\theta(x)$, let $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ be the sequence of likelihood ratios at sample size n, $n\in\Gamma^+$. Then

$$\lim_{n\to\infty} E_{\theta_0} L_{n,\theta_0,\theta_1}^{\frac{1}{2}} = 0 \quad \text{iff} \quad P_{\theta_0}(L_{1,\theta_0,\theta_1} = 1) < 1 \qquad (1.19)$$

and

$$\lim_{n\to\infty} E_{\theta_1} L_{n,\theta_0,\theta_1}^{-\frac{1}{2}} = 0 \quad \text{iff} \quad P_{\theta_1}(L_{1,\theta_0,\theta_1}^{-1} = 1) < 1 \qquad (1.20),$$

respectively.

P r o o f. Because the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables, we obtain

$$L_{n,\theta_0,\theta_1} = \prod_{i=1}^{n} (f_{\theta_1}(X_i)/f_{\theta_0}(X_i)) \ , \ n\in\Gamma^+ .$$

This implies

$$E_{\theta_0} L_{n,\theta_0,\theta_1}^{\frac{1}{2}} = E_{\theta_0}(\prod_{i=1}^{n}(f_{\theta_1}(X_i)/f_{\theta_0}(X_i))^{\frac{1}{2}}$$

$$= \prod_{i=1}^{n} E_{\theta_0}(f_{\theta_1}(X_i)/f_{\theta_0}(X_i))^{\frac{1}{2}} = (E_{\theta_0} L_{1,\theta_0,\theta_1}^{\frac{1}{2}})^n .$$

Hence, we have

$$\lim_{n\to\infty} E_{\theta_0} L_{1,\theta_0,\theta_1}^{\frac{1}{2}} = 0 \quad \text{iff} \quad E_{\theta_0} L_{1,\theta_0,\theta_1}^{\frac{1}{2}} < 1 . \qquad (1.21)$$

Applying Schwarz's inequality we obtain

$$E_{\theta_0} L_{1,\theta_0,\theta_1}^{\frac{1}{2}} = E_{\theta_0}(L_{1,\theta_0,\theta_1}^{\frac{1}{2}} \cdot 1) \leqslant (E_{\theta_0} L_{1,\theta_0,\theta_1})^{\frac{1}{2}} 1 = 1, \qquad (1.21')$$

where the strict equality holds iff a real number c exists with

$$P_{\theta_0}(L_{1,\theta_0,\theta_1}^{\frac{1}{2}} = c) = 1 .$$

This condition implies $P_{\theta_0}(L_{1,\theta_0,\theta_1} = c^2) = 1$ and $E_{\theta_0} L_{1,\theta_0,\theta_1} = c^2$. Otherwise, we have $E_{\theta_0} L_{1,\theta_0,\theta_1} = 1$. Hence we obtain $c = 1$. Thus,

the strict inequality holds in (1.22) iff $P_{\theta_o}(L_{1,\theta_o,\theta_1} = 1) < 1$. This, together with (1.21), proves (1.19). Analogously, we can establish (1.20). ■

We remark, that the conditions

$$P_{\theta_o}(L_{1,\theta_o,\theta_1} = 1) < 1 \quad \text{and} \quad P_{\theta_1}(L^{-1}_{1,\theta_o,\theta_1} = 1) < 1$$

of Lemma 1.5.2 may also be written as

$$P_{\theta_o}(f_{\theta_o}(X_1) = f_{\theta_1}(X_1)) < 1 \text{ and } P_{\theta_1}(f_{\theta_o}(X_1) = f_{\theta_1}(X_1)) < 1,$$

respectively, which emphasizes that these conditions (1.21) are very weak. We note that the considered convergence properties will also play a role in connection with the termination property of a test.

1. 6 Conjugated parameter pairs

Frequently, the parameter space Θ contains more than two parameters. If we consider in such a situation a test (N, δ) for a simple hypothesis $H_o: \theta = \theta_o$ against a simple alternative hypothesis $H_1: \theta = \theta_1$, $\theta_o \neq \theta_1$, then we are also interested in properties of this test if the true parameter θ does not coincide with θ_o or θ_1. In these cases, conjugated parameter pairs are a helpful tool. For the first time certain aspects of such parameter pairs were considered by GIRSHICK [36] in connection with the approximate computation of the operating characteristic function and the ASN-function of WALD's sequential likelihood ratio test. later on by BLASBALG [16], LECHNER [52] BERK [13] and SAVAGE [66]. The starting point of their conjugacy concept is the fact that several parameter pairs θ_o and θ_1 and risks α and β can lead to the same WALD's likelihood ratio test, that means that they involve in their conjugacy concept the special structure of WALD's likelihood ratio test.

Here we introduce a conjugacy concept allowing us to obtain relations between two parameter pairs of parameter space Θ without any reference to the explicit structure of the underlying test, considering certain properties of the likelihood ratio sequence. This concept will be a helpful tool for the whole theory of sequential tests.

For all further investigations we will assume that the considered statistical structure $(\Omega, \mathfrak{F}, \mathfrak{P})$ is characterized as follows. The family $\mathfrak{P}$ of probability measures is indexed by a parameter $\theta \in \Theta$

so that $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$. Further, a sequence $\{\mathcal{F}_n\}_{n \in \Gamma^+}$ of non-decreasing sub-σ-algebras of $\mathcal{F}$ is given, where for every pair $\theta', \theta'' \in \Theta$ and $n \in \Gamma^+$ the corresponding restrictions $P^{(n)}_{\theta'}$ and $P^{(n)}_{\theta''}$ of $P_{\theta'}$ and $P_{\theta''}$, respectively, from $(\Omega, \mathcal{F})$ to $(\Omega, \mathcal{F}_n)$ are equivalent measures. Then the Radon-Nikodym-derivatives $dP^{(n)}_{\theta''}/dP^{(n)}_{\theta'}$ of $P^{(n)}_{\theta''}$ w.r.t $P^{(n)}_{\theta'}$ and $dP^{(n)}_{\theta'}/dP^{(n)}_{\theta''}$ of $P^{(n)}_{\theta'}$ w.r.t. $P^{(n)}_{\theta''}$ exist for every $n \in \Gamma^+$, where $dP^{(n)}_{\theta'}/dP^{(n)}_{\theta''} = (dP^{(n)}_{\theta''}/dP^{(n)}_{\theta'})^{-1}$. We shall again denote any versions of these derivatives by $L_{n,\theta',\theta''}$ and $L_{n,\theta'',\theta'}$, respectively. For this situation we define the following.

<u>Definition 1.6.1</u>. Two parameter pairs $\theta', \theta'' \in \Theta$, $\theta' \neq \theta''$, and $\theta_0, \theta_1 \in \Theta$, $\theta_0 \neq \theta_1$, are said to be <u>conjugated</u> iff a real number $h \neq 0$ exists so that for every $n \in \Gamma^+$

$$L_{n,\theta',\theta''} = L^h_{n,\theta_0,\theta_1}, \quad \omega \in \Omega . \tag{1.22}$$

We shall write then: $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$.

We remark, that it would be possible to substitute the condition (1.22) in Definition 1.6.1 by the weaker condition

$$P^{(n)}_{\theta'}(L_{n,\theta',\theta''} = L^h_{n,\theta_0,\theta_1}) = 1,$$

but in view of a more comprehensible representation we renounce this generalization.

Some conclusions following immediately from the above definition:

(i) If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$, then $(\theta'',\theta') \overset{-h}{\sim} (\theta_0,\theta_1)$.

Especially we have

$$(\theta_0,\theta_1) \overset{1}{\sim} (\theta_0,\theta_1) \quad \text{and} \quad (\theta_1,\theta_0) \overset{-1}{\sim} (\theta_0,\theta_1).$$

(ii) If $(\theta',\theta'') \overset{h_1}{\sim} (\hat\theta,\hat{\hat\theta})$ and $(\hat\theta,\hat{\hat\theta}) \overset{h_2}{\sim} (\theta_0,\theta_1)$, then $(\theta',\theta'') \overset{h_1 h_2}{\sim} (\theta_0,\theta_1)$.

(iii) If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$, then $(\theta_0,\theta_1) \overset{h^{-1}}{\sim} (\theta',\theta'')$.

(iv) If $(\theta',\theta'') \overset{h_1}{\sim} (\theta_0,\theta_1)$ and $(\hat\theta,\hat{\hat\theta}) \overset{h_2}{\sim} (\theta_0,\theta_1)$, then $(\theta',\theta'') \overset{h_1 h_2^{-1}}{\sim} (\hat\theta,\hat{\hat\theta})$.

For any given sample size N w.r.t. $\{\mathcal{F}_n\}_{n \in \Gamma^+}$ and any real h let $L^h_{N,\theta_0,\theta_1}$ be defined by

$$L^h_{N,\theta_0,\theta_1} = \sum_{n \in \bar\Gamma^+} L^h_{n,\theta_0,\theta_1} \chi_{\{N=n\}},$$

where

$$L^h_{\infty,\theta_0,\theta_1} = \limsup_{n \to \infty} L^h_{n,\theta_0,\theta_1} .$$

Then $L^h_{N,\theta_o,\theta_1}$ is an $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable random variable (see [40], Theorem 1.5) and the following transformation rule holds for expectation values of $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable random variables in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$.

L e m m a 1.6.1. Let N be a sample size w.r.t. $\{\mathfrak{F}_n\}_{n\in\Gamma^+}$, let $\varphi:\Omega \to R^1$ be an $(\mathfrak{F}_N, \mathfrak{B}^1)$- measurable and $P_{\theta''}$-integrable function. If $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$, then for every $F \in \mathfrak{F}_N$

$$E_{\theta'}(\varphi L^h_{N,\theta_o,\theta_1} \chi_{\{N<\infty\}} | F)P_{\theta'}(F) = E_{\theta''}(\varphi \chi_{\{N<\infty\}} | F)P_{\theta''}(F). \tag{1.23}$$

P r o o f. We have

$$\begin{aligned} E_{\theta''}(\varphi\chi_{\{N<\infty\}} | F) &= E_{\theta''}(\varphi\chi_{\{N<\infty\}} | F)E_{\theta''}\chi_F \\ &= E_{\theta''}(\varphi\chi_{\{N<\infty\}} | F)P_{\theta''}(F). \end{aligned} \tag{1.24}$$

Since $F \in \mathfrak{F}_N$ by [40], Lemma 1.4, a sequence $\{F_n\}_{n\in\bar\Gamma^+}$ of sets $F_n \in \mathfrak{F}_n$, exists with

$$F = \sum_{n\in\bar\Gamma^+} \chi_{F_n} \chi_{\{N=n\}}.$$

Further, since φ is an $(\mathfrak{F}_N, \mathfrak{B}^1)$-measurable function a sequence $\{\varphi_n\}_{n\in\bar\Gamma^+}$ of $(\mathfrak{F}_n, \mathfrak{B}^1)$-measurable functions $\varphi_n:\Omega \to R^1$, $n\in\bar\Gamma^+$, exists with

$$\varphi = \sum_{n\in\bar\Gamma^+} \varphi_n \chi_{\{N=n\}}.$$

Hence, by $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ we obtain

$$\begin{aligned} E_{\theta''}(\varphi\chi_{\{N<\infty\}}\chi_F) &= \sum_{n\in\Gamma^+} \int_{\{N=n\}} \varphi\chi_F dP_{\theta''} \\ &= \sum_{n\in\Gamma^+} \int_{\{N=n\}} \varphi_n \chi_{F_n} dP^{(n)}_{\theta''} \\ &= \sum_{n\in\Gamma^+} \int_{\{N=n\}} \varphi_n \chi_{F_n} L^h_{n,\theta_o,\theta_1} dP^{(n)}_{\theta'} \\ &= \sum_{n\in\Gamma^+} \int_{\{N=n\}} \varphi L^h_{N,\theta_o,\theta_1} \chi_F dP_{\theta'} \\ &= E_{\theta'}(\varphi L^h_{N,\theta_o,\theta_1} \chi_{\{N<\infty\}} \chi_F) \end{aligned}$$

$$= E_{\theta'}(\varphi L^h_{N,\theta_o,\theta_1} \chi_{\{N<\infty\}} | F) P_{\theta'}(F).$$

This together with (1.24) implies (1.23). ■

The identity (1.23) will be a useful device in connection with the investigation into the properties of sequential tests. Two special cases of (1.23) will be of particular interest.

(i) If $P_{\theta'}(N<\infty) = 1$ and $P_{\theta'}(F) > 0$, then (1.23) with $\varphi = 1$ implies

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1} | F) = P_{\theta''}(F)/P_{\theta'}(F) . \tag{1.25}$$

(ii) If $P_{\theta'}(N<\infty) = 1$ and $F = \Omega$, then (1.23) implies

$$E_{\theta''}\varphi = E_{\theta'}\varphi L^h_{N,\theta_o,\theta_1} . \tag{1.26}$$

The subsequent lemma presents a conjugacy criterion for the i.i.d.-case considered in the Examples 1.2.1 and 1.4.1. We suppose that P_θ, $\mathcal{F}_n$ and $P^{(n)}_\theta$ are determined like in these examples.

<u>L e m m a 1.6.2</u>. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having a density $f_\theta(x)$, where the set $S = \{x \in \mathfrak{X}: f_\theta(x) > 0\}$ does not depend on $\theta \in \Theta$. If to any given $\theta', \theta'', \theta_o, \theta_1 \in \Theta$, $\theta' \neq \theta''$, $\theta_o \neq \theta_1$, a real number $h \neq 0$ exists so that

$$f_{\theta''}(x)/f_{\theta'}(x) = (f_{\theta_1}(x)/f_{\theta_o}(x))^h \quad \text{for} \quad x \in S, \tag{1.27}$$

then we have $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$.

P r o o f. Under the conditions of this lemma we obtain for every pair $\hat{\theta}, \hat{\hat{\theta}} \in \Theta$ and $n \in \Gamma^+$ a version $L_{n,\hat{\theta},\hat{\hat{\theta}}}$ of $dP^{(n)}_{\hat{\hat{\theta}}}/dP^{(n)}_{\hat{\theta}}$ if we put

$$L_{n,\hat{\theta},\hat{\hat{\theta}}}(\omega) = \prod_{i=1}^{n} (f_{\hat{\hat{\theta}}}(X_i(\omega))/f_{\hat{\theta}}(X_i(\omega))) \ , \ \omega \in \Omega , \tag{1.28}$$

where in case of $f_{\hat{\theta}}(X_i(\omega)) = 0$ we can choose an arbitrary value for $L_{n,\hat{\theta},\hat{\hat{\theta}}}$. We consider the set $\Omega_S = \{\omega \in \Omega: f_\theta(X_i(\omega)) > 0\}$, $\theta \in \Theta$, $i \in \Gamma^+$. By definition of S and the i.i.d.-property of $\{X_n\}_{n\in\Gamma^+}$ we obtain $\Omega_S = \{\omega \in \Omega : X_i(\omega) \in S\}$ so that Ω_S does not depend on $\theta \in \Theta$ and $i \in \Gamma^+$. Hence, for $\omega \in \Omega_S$ we obtain by (1.28) and (1.27) for every $n \in \Gamma^+$

$$L_{n,\theta',\theta''}(\omega) = \prod_{i=1}^{n} (f_{\theta''}(X_i(\omega))/f_{\theta'}(X_i(\omega)))$$

$$= \prod_{i=1}^{n} (f_{\theta_1}(X_i(\omega))/f_{\theta_0}(X_i(\omega)))^h$$

$$= L^h_{n,\theta_0,\theta_1}(\omega) \; . \tag{1.29}$$

For $\omega \in \Omega - \Omega_S$ $L_{n,\theta',\theta''}$ and L_{n,θ_0,θ_1} can be defined in an arbitrary manner. Thus, we achieve that (1.29) holds also for $\omega \in \Omega - \Omega_S$. Hence, we obtain (1.22) and the proof is complete.■

This lemma can be modified for exponential families as follows.

<u>Lemma 1.6.3</u>. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having the density

$$f_\theta(x) = h(x)\exp(d(\theta)t(x) - c(\theta)) \; , \quad x \in \mathcal{X} \; , \; \theta \in \Theta \; . \tag{1.30}$$

Suppose that $t(X_1)$ has a non-degenerated distribution w.r.t. P_θ for every $\theta \in \Theta$. Then we have $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ iff

$$d(\theta'') - d(\theta') = h(d(\theta_1) - d(\theta_0)) \tag{1.31}$$

and

$$c(\theta'') - c(\theta') = h(c(\theta_1) - c(\theta_0)) \; . \tag{1.32}$$

P r o o f. We have $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ iff versions $L_{n,\theta',\theta''}$ and L_{n,θ_0,θ_1} of the Radon-Nikodym-derivatives $dP^{(n)}_{\theta''}/dP^{(n)}_{\theta'}$ and $dP^{(n)}_{\theta_1}/dP^{(n)}_{\theta_0}$ exist for every $n \in \Gamma^+$ so that $L_{n,\theta',\theta''} = L^h_{n,\theta_0,\theta_1}$ holds for a non-zero h. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables with a density $f_\theta(x)$ we can choose $L_{n,\theta',\theta''}$ and L_{n,θ_0,θ_1} for every $n \in \Gamma^+$ according to

$$L_{n,\theta',\theta''} = \prod_{i=1}^{n} (f_{\theta''}(X_i)/f_{\theta'}(X_i)) \quad \text{and}$$

$$L_{n,\theta_0,\theta_1} = \prod_{i=1}^{n} (f_{\theta_1}(X_i)/f_{\theta_0}(X_i)) \; .$$

Then by (1.30) we have $L_{n,\theta',\theta''} = L^h_{n,\theta_0,\theta_1}$, $n \in \Gamma^+$, iff

$$\left.\begin{aligned} &\underbrace{(d(\theta'') - d(\theta'))}\sum_{i=1}^{n} t(X_i) - \underbrace{(c(\theta'') - c(\theta'))} = \\ &\underbrace{h(d(\theta_1) - d(\theta_0))}\sum_{i=1}^{n} t(X_i) - \underbrace{h(c(\theta_1) - c(\theta_0))} \; , \; n \in \Gamma^+ . \end{aligned}\right\} *_1$$

$*_1$ → in the underlined part the author is using equation 1·31 and in the underlined part the author is using equation 1·32.

Since $t(X_1)$ has a non-degenerated distribution w.r.t P_θ for $\theta \in \Theta$, this is true iff (1.31) and (1.32) hold and the proof is complete. ■

We notice, if (1.31) und (1.32) hold, then we have further

$$\frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)} = \frac{c(\theta'') - c(\theta')}{d(\theta'') - d(\theta')} \,. \tag{1.33}$$

This equation, together with (1.31) and (1.32), can be used to determinate to given $\theta', \theta_o, \theta_1 \in \Theta$ a $\theta'' \in \Theta$ and an h so that $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$. Tasks of this typ will arise, e.g., in connection with an approximate computation of the power function of WALD's likelihood ratio test. We refer to Section 2. We remark that the system of equations (1.31) and (1.32) to given θ' as a system of equations in the unknowns θ'' and h has the trivial solution $\theta'' = \theta'$ and $h = 0$ in each case. Indeed, this trivial solution will be of importance only if it is the only solution to the considered system of equations.

<u>E x a m p l e 1.6.1</u>. We consider some special cases of (1.30).
(i) <u>The Bernoulli distribution</u>. We have

$$f_\theta(x) = \theta^x (1 - \theta)^{1-x} \;,\; x \in \{0,1\} \;,\; \theta \in (0,1) \;.$$

This implies

$$c(\theta) = -\ln(1 - \theta) \quad \text{and} \quad d(\theta) = \ln(\theta/(1-\theta)) \;,$$

and we obtain the system of equations

$$\ln \frac{\theta''(1-\theta')}{\theta'(1-\theta'')} = h \ln \frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)} \tag{1.34}$$

and

$$\ln \frac{1-\theta''}{1-\theta'} = h \ln \frac{1-\theta_1}{1-\theta_o} \;, \tag{1.35}$$

according to (1.31) and (1.32). In general, it is not possible to find to given $\theta' \in \Theta$ an explicit solution for h and θ''. Indeed, (1.34) and (1.35) implies

$$h = \ln \frac{1-\theta'(\theta_1/\theta_o)^h}{1-\theta'} \Big/ \ln \frac{1-\theta_1}{1-\theta_o} \tag{1.36}$$

and this equation can be used to obtain to given $\theta' \in \Theta$ a solution $h \neq 0$ by the method of iteration if

$$\theta' \neq \theta^* = \ln \frac{1-\theta_o}{1-\theta_1} \Big/ \ln \frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)} \;. \tag{1.37}$$

Then, (1.35) implies

$$\theta'' = 1 - (1-\theta')(\frac{1-\theta_1}{1-\theta_o})^h . \tag{1.38}$$

Conversely, it is not difficult to obtain to given $h \neq 0$ a parameter pair (θ',θ'') with $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$. It follows from (1.34) and (1.35) if $h \neq 0$

$$\theta' = (1-(\frac{1-\theta_o}{1-\theta_1})^h)/(1-(\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)})^h) \tag{1.39}$$

and θ'' can be determined again by (1.38). An explicit solution to the system of equations (1.31) and (1.32) can be obtained if

$$\theta_o = \frac{1}{2} - \varepsilon \quad \text{and} \quad \theta_1 = \frac{1}{2} + \varepsilon \qquad \text{with } \varepsilon \in (0,\tfrac{1}{2}).$$

Then we obtain by (1.34) and (1.35)

$$\theta'' = 1 - \theta' \tag{1.40}$$

and

$$h = \ln \frac{1-\theta'}{\theta'} \Big/ \ln \frac{1+2\varepsilon}{1-2\varepsilon} \tag{1.41}$$

for $\theta' \in (0,1)$ and $\theta' \neq \theta^* = 1/2$.

(ii) <u>The Poisson distribution</u>. We have

$$f_\theta(x) = \frac{\theta^x}{x!}\exp(-\theta x) , \quad x \in \Gamma_o^+ , \ \theta \in (0, \infty) = \Theta$$

and we obtain

$$c(\theta) = \theta \quad \text{and} \quad d(\theta) = \ln \theta .$$

Hence, by (1.31) and (1.32) we obtain the system of equations

$$\ln \theta'' - \ln \theta' = h(\ln \theta_1 - \ln \theta_o) \tag{1.42}$$

and

$$\theta'' - \theta' = h(\theta_1 - \theta_o). \tag{1.43}$$

It is not possible to find to given θ' an explicit solution for h and θ'', but (1.42) and (1.43) implies

$$h = \theta'((\theta_1/\theta_o)^h - 1)/(\theta_1 - \theta_o) . \tag{1.44}$$

This equation can be used to obtain to given $\theta' \in \Theta$ a solution h by the method of iteration if

$$\theta' \neq \theta^* = (\theta_1 - \theta_o)/\ln(\theta_1/\theta_o) . \tag{1.45}$$

Then (1.43) implies

$$\theta'' = \theta' + h(\theta_1 - \theta_o) \; . \tag{1.46}$$

Otherwise, we can obtain for every non-zero h a parameter pair (θ',θ'') with $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$. Then (1.42) and (1.43) implies

$$\theta' = h(\theta_1 - \theta_o)/((\theta_1/\theta_o)^h - 1) \tag{1.47}$$

and θ'' can be determined by (1.46).

(iii) <u>The normal distribution</u>. We consider the mean and suppose that the variance σ^2 is known. Then we have

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-\theta)^2}{2\sigma^2}) \; , \; x \in R^1 \; , \; \theta \in R^1 = \Theta$$

with

$$c(\theta) = \theta^2/2\sigma^2 \quad \text{and} \quad d(\theta) = \theta/\sigma^2 \; .$$

According to (1.31) and (1.32), we obtain

$$\theta'' - \theta' = h(\theta_1 - \theta_o) \tag{1.48}$$

and

$$\theta''^2 - \theta'^2 = h(\theta_1^2 - \theta_o^2) \; . \tag{1.49}$$

This example is one of the few examples, where to given θ' the system of equations (1.31) and (1.32) can be solved explicitly. The corresponding solution is

$$h = (\theta_o + \theta_1 - 2\theta')/(\theta_1 - \theta_o) \; , \tag{1.50}$$

$$\theta'' = h(\theta_1 - \theta_o) + \theta' \; . \tag{1.51}$$

We note, that this solution does not depend on σ^2 and we have a non-zero solution for h iff

$$\theta' \neq \theta^* = (\theta_o + \theta_1)/2 \; . \tag{1.52}$$

Further, (1.50), (1.51) and (1.52) implies

$$\theta'' = 2\theta \;\; - \theta' \; . \tag{1.53}$$

(iv) <u>The exponential distribution</u>. We suppose

$$f_\theta(x) = \theta \exp(-\theta x) \; , \quad x \in (0,\infty) \; , \quad \theta \in (0,\infty) = \Theta.$$

Then we have

$$c(\theta) = -\ln\theta \quad \text{and} \quad d(\theta) = -\theta \; .$$

According to (1.31) and (1.32), we obtain the system of equations

$$-\theta'' + \theta' = h(-\theta_1 + \theta_0),$$

$$-\ln \theta'' + \ln \theta' = h(-\ln \theta_1 + \ln \theta_0).$$

This system of equations is identical with the system of equations (1.42) and (1.43) obtained for the Poisson distribution and here we obtain the same relations between h, θ', θ'', θ_0 and θ_1 like in the Poisson case. ■

Now we consider some further relations between $\theta',\theta'',\theta_0,\theta_1$ and h in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ for the one-parametric exponential family.

L e m m a 1.6.4. Let X_1 be a random variable with the density

$$f_\theta(x) = h(x)\exp(d(\theta)t(x) - c(\theta)) \quad , \; x \in \mathfrak{X} \quad , \; \theta \in (\underline{\theta},\overline{\theta}), \tag{1.54}$$

where c and d are analytic functions on $(\underline{\theta},\overline{\theta})$, d is strictly monotonical in θ on $(\underline{\theta},\overline{\theta})$ and $D^2_\theta t(X_1) > 0$ for $\theta \in (\underline{\theta},\overline{\theta})$. For $\theta,\hat{\theta} \in (\underline{\theta},\overline{\theta})$ let $\zeta(\theta,\hat{\theta})$ be defined by

$$\zeta(\theta,\hat{\theta}) = (c(\theta) - c(\hat{\theta}))d'(\hat{\theta}) - (d(\theta) - d(\hat{\theta}))c'(\hat{\theta}). \tag{1.55}$$

We suppose

$$\lim_{\theta \to \underline{\theta}} \zeta(\theta,\hat{\theta}) = \lim_{\theta \to \overline{\theta}} \zeta(\theta,\hat{\theta}) = \infty. \tag{1.56}$$

Then we have:

(i) To each $\theta' < \hat{\theta}$ corresponds a $\theta'' > \hat{\theta}$ so that

$$\zeta(\theta',\hat{\theta}) = \zeta(\theta'',\hat{\theta}) > 0. \tag{1.57}$$

This correspondence is a one-to-one correspondence between the elements of $(\underline{\theta},\hat{\theta})$ and the elements of $(\hat{\theta},\overline{\theta})$, respectively.

(ii) For every pair (θ_0,θ_1), $\underline{\theta} < \theta_0 < \theta_1 < \overline{\theta}$, a uniquely determined parameter θ^*, $\theta_0 < \theta^* < \theta_1$, exists, so that

$$\frac{c'(\theta^*)}{d'(\theta^*)} = \frac{c(\theta_1) - c(\theta_0)}{d(\theta_1) - d(\theta_0)} \tag{1.58}$$

and

$$\zeta(\theta_0,\theta^*) = \zeta(\theta_1,\theta^*) \tag{1.59}$$

holds.

P r o o f. (i) We note that for the distribution family (1.54)

$$E_\theta t(X_1) = \frac{c'(\theta)}{d'(\theta)} \tag{1.60}$$

and

$$D^2_\theta t(X_1) = \frac{d}{d\theta} E_\theta t(X_1) \tag{1.61}$$

holds (see e.g. LEHMANN [53]). Since $D_\theta^2 t(X_1) > 0$ for $\theta \in (\underline{\theta},\bar{\theta})$, we obtain by (1.60) and (1.61) $c'(\theta)/d'(\theta)$ is strictly monotonically increasing in θ on $(\underline{\theta},\bar{\theta})$. Thus we have

$$\frac{c'(\theta)}{d'(\theta)} < \frac{c'(\hat{\theta})}{d'(\hat{\theta})} \quad \text{for} \quad \underline{\theta} < \theta < \hat{\theta} \tag{1.62}$$

and

$$\frac{c'(\theta)}{d'(\theta)} > \frac{c'(\hat{\theta})}{d'(\hat{\theta})} \quad \text{for} \quad \hat{\theta} < \theta < \bar{\theta}. \tag{1.63}$$

Now we consider the first derivative of $\zeta(\theta,\hat{\theta})$ w.r.t. θ and obtain

$$\zeta'(\theta,\hat{\theta}) = c'(\theta)d'(\hat{\theta}) - d'(\theta)c'(\hat{\theta}).$$

Then we have $\zeta'(\hat{\theta},\hat{\theta}) = 0$ and the inequality (1.62) implies

$$\zeta'(\theta,\hat{\theta}) = d'(\theta)d'(\hat{\theta})\left(\frac{c'(\theta)}{d'(\theta)} - \frac{c'(\hat{\theta})}{d'(\hat{\theta})}\right) < 0 \text{ for } \underline{\theta} < \theta < \hat{\theta}.$$

That means that $\zeta(\theta,\hat{\theta})$ is strictly monotonically decreasing in θ on $(\underline{\theta},\hat{\theta})$. Analogously, we obtain $\zeta(\theta,\hat{\theta})$ is strictly monotonically increasing in θ on $(\hat{\theta},\bar{\theta})$. Hence, $\zeta(\theta,\hat{\theta})$ as a function of θ has one and only one minimum at $\theta = \hat{\theta}$. Then, by $\zeta(\hat{\theta},\hat{\theta}) = 0$ we obtain $\zeta(\theta,\hat{\theta}) > 0$ for every $\theta \neq \hat{\theta}$. Furtheron, the monotonicity properties of $\zeta(\theta,\hat{\theta})$, the fact $\zeta(\hat{\theta},\hat{\theta}) = 0$ and (1.56) imply, for every ζ_0, $0 < \zeta_0 < +\infty$, there exists a uniquely determined pair θ', θ'' with $\underline{\theta} < \theta' < \hat{\theta} < \theta'' < \bar{\theta}$ and $\zeta(\theta',\hat{\theta}) = \zeta(\theta'',\hat{\theta}) = \zeta_0$. If ζ_0 ranges in $(0,\infty)$ we obtain the proposed one-to-one correspondence between the elements of $(\underline{\theta},\hat{\theta})$ and the elements of $(\hat{\theta},\bar{\theta})$.

(ii) By the Cauchy Theorem, a parameter θ^* exists for every pair θ_0,θ_1, $\underline{\theta} < \theta_0 < \theta_1 < \bar{\theta}$, where $\theta_0 < \theta^* < \theta_1$ and (1.58) holds. Since $c'(\theta)/d'(\theta)$ is strictly monotonically increasing in θ on $(\underline{\theta},\bar{\theta})$, θ^* is uniquely determined. By definition of ζ, the equations (1.58) and (1.59) are equivalent for $\hat{\theta} = \theta^*$. This completes the proof.■

By means of this lemma we further obtain the following conjugacy criterion.

<u>Lemma 1.6.5</u>. Let $\{X_n\}_{n \in \mathbb{N}^+}$ be a sequence of i.i.d. random variables with the density (1.54). Suppose that Lemma 1.6.4 holds. If to a given parameter pair θ_0,θ_1, $\underline{\theta} < \theta_0 < \theta_1 < \theta$, the parameter θ^* is determined by (1.58), then we have $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ iff

$$\zeta(\theta',\theta^*) = \zeta(\theta'',\theta^*) > 0 \tag{1.64}$$

and

$$h = \frac{d(\theta'') - d(\theta')}{d(\theta_1) - d(\theta_0)} \,. \tag{1.65}$$

Especially we obtain

$$h > 0 \quad \text{for} \quad \underline{\theta} < \theta' < \theta^* \tag{1.66}$$

and

$$h < 0 \quad \text{for} \quad \theta^* < \theta' < \overline{\theta} \tag{1.67}$$

P r o o f. The equation (1.64) is equivalent to

$$\frac{c'(\theta^*)}{d'(\theta^*)} = \frac{c(\theta'') - c(\theta')}{d(\theta'') - d(\theta')} \tag{1.68}$$

and $\xi(\theta',\theta^*) > 0$ and $\xi(\theta'',\theta^*) > 0$ implies $\theta' \neq \theta^*$ and $\theta'' \neq \theta^*$. Furtheron, since θ^* is determined by (1.58), we have also

$$\frac{c'(\theta^*)}{d'(\theta^*)} = \frac{c(\theta_1) - c(\theta_0)}{d(\theta_1) - d(\theta_0)} \,. \tag{1.69}$$

If now h is determined by (1.65) then we obtain

$$d(\theta'') - d(\theta') = h(d(\theta_1) - d(\theta_0)). \tag{1.70}$$

This, together with (1.68) and (1.69), provides

$$c(\theta'') - c(\theta') = h(c(\theta_1) - c(\theta_0)). \tag{1.71}$$

Applying Lemma 1.6.3 we obtain $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$.
Conversely, by Lemma 1.6.3 $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ implies (1.70) and (1.71) and h is determined by (1.70) so that (1.65) holds. Further, (1.70), (1.71) and (1.58) implies (1.68) which is equivalent to $\xi(\theta',\theta^*) = \xi(\theta'',\theta^*)$. Since $h \neq 0$, we obtain $\theta' \neq \theta''$, and therefore $\xi(\theta',\theta^*) = \xi(\theta'',\theta^*) > 0$ so that also (1.64) holds.
In order to establish (1.66) and (1.67) we note that, according to Lemma 1.6.4 (ii) $\theta_0 < \theta^* < \theta_1$ holds. Further, if $\underline{\theta} < \theta' < \theta^*$, then the equation (1.64) and Lemma 1.6.4 (i) imply $\theta'' > \theta^*$. Thus, since d is strictly monotonically increasing, we obtain by (1.65) $h > 0$ for $\underline{\theta} < \theta' < \theta^*$. Analogously, we obtain (1.67).■

We discuss some consequences of Lemma 1.6.4 and Lemma 1.6.5:

(i) If for the family (1.54) to a given parameter pair θ_0,θ_1, $\underline{\theta} < \theta_0 < \theta_1 < \overline{\theta}$, the parameter θ^* is determined by (1.58) and if $\{X_n\}_{n\in\Gamma^+}$ forms a sequence of i.i.d. random variables, then the logarithm of the likelihood ratio $Z_{n,\theta_0,\theta_1} = \ln L_{n,\theta_0,\theta_1}$ can be written as

$$Z_{n,\theta_0,\theta_1} = (d(\theta_1) - d(\theta_0)) \sum_{i=1}^{n} t(X_i) - n(c(\theta_1) - c(\theta_0))$$

$$= d_{\theta_0,\theta_1}\Big(\sum_{i=1}^{n} t(X_i) - n\frac{c'(\theta^*)}{d'(\theta^*)} \Big) \,, \; n\in\Gamma^+ \tag{1.72}$$

with

$$d_{\theta_o,\theta_1} = d(\theta_1) - d(\theta_o). \tag{1.73}$$

Since $E_\theta t(X_i) = c'(\theta)/d'(\theta)$ for $i \in \Gamma^+$, we have instead of (1.72) also

$$Z_{n,\theta_o,\theta_1} = d_{\theta_o,\theta_1}\left(\sum_{i=1}^{n} t(X_i) - nE_{\theta^*}t(X_1)\right). \tag{1.74}$$

If moreover $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ holds, then we obtain by (1.33), (1.65), (1.73) and (1.72)

$$\begin{aligned} Z_{n,\theta',\theta''} &= (d(\theta'') - d(\theta'))\sum_{i=1}^{n} t(X_i) - n(c(\theta'') - c(\theta')) \\ &= hd_{\theta_o,\theta_1}\left(\sum_{i=1}^{n} t(X_i) - n\frac{c'(\theta^*)}{d'(\theta^*)}\right) \\ &= hd_{\theta_o,\theta_1}\left(\sum_{i=1}^{n} t(X_i) - nE_{\theta^*}t(X_1)\right),\ n \in \Gamma^+, \end{aligned} \tag{1.75}$$

so that in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ the logarithm of likelihood ratio $Z_{n,\theta',\theta''}$ is always proportional to

$$\sum_{i=1}^{n} t(X_i) - n\,\frac{c'(\theta^*)}{d'(\theta^*)} = \sum_{i=1}^{n} t(X_i) - nE_{\theta^*}t(X_1). \tag{1.76}$$

If especially $t(X_1) = X_1$ and $E_\theta X_1 = \theta$, then we obtain in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ even

$$Z_{n,\theta',\theta''} = hd_{\theta_o,\theta_1}\left(\sum_{i=1}^{n} X_i - n\theta^*\right),\ n \in \Gamma^+, \tag{1.77}$$

and

$$\theta^* = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)}. \tag{1.78}$$

(ii) Parameter θ^* introduced by the second part of Lemma 1.6.4 can be used to obtain a partition of the parameter space $\Theta = (\underline{\theta},\overline{\theta})$, according to

$$\Theta = (\underline{\theta},\theta^*) \cup \{\theta^*\} \cup (\theta^*,\overline{\theta}),$$

where in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ (1.66) and (1.67) holds. Partitions of this kind may be of interest in connection with tests for the hypothesis

$$H_o\colon \theta \leq \theta^* \quad \text{against} \quad H_1\colon \theta > \theta^*, \quad \theta^* \in (\underline{\theta},\overline{\theta}),$$

where in a sense parameter θ^* can be interpreted as a <u>separating-parameter</u>. For, if instead of parameters θ_o and θ_1 this separating-

parameter θ^* is given, the by Lemma 1.6.4 a continuum of pairs θ', θ'', $\underline{\theta} < \theta' < \theta^* < \theta'' < \tilde{\theta}$, exists so that $\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*)$ holds. Hence, by Lemma 1.6.5 and according to our above remark the logarithm of the likelihood ratio $Z_{n,\theta',\theta''}$ is for every pair θ', θ'' satisfying $\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*)$ proportional to

$$\sum_{i=1}^{n} t(X_i) - n\frac{c'(\theta^*)}{d'(\theta^*)} = \sum_{i=1}^{n} t(X_i) - nE_{\theta^*}t(X_1).$$

This property will play a role in connection with so-called locally most powerful tests considered in Lemma 2.2.1.

E x a m p l e 1.6.2. We consider the separating-parameters for the distributions of Example 1.6.1.

(i) The Bernoulli distribution. We have $E_\theta t(X_1) = E_\theta X_1 = \theta$, $c(\theta) = -\ln(1-\theta)$ and $d(\theta) = \ln(\theta/(1-\theta))$. According to (1.78), this implies for $0 < \theta_o, \theta_1 < 1$, $\theta_o \neq \theta_1$,

$$\theta^* = \ln\frac{1-\theta_o}{1-\theta_1} \Big/ \ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)} . \qquad (1.79)$$

If $\theta_o = \frac{1}{2} - \varepsilon$ and $\theta_1 = \frac{1}{2} + \varepsilon$, $0 < \varepsilon < \frac{1}{2}$, then we obtain $\theta^* = \frac{1}{2}$.

(ii) The Poisson distribution. We have $E_\theta t(X_1) = E_\theta X_1 = \theta$, $c(\theta) = \theta$ and $d(\theta) = \ln\theta$. Then (1.78) implies for $0 < \theta_o, \theta_1 < \infty$, $\theta_o \neq \theta_1$,

$$\theta^* = (\theta_1 - \theta_o)/\ln(\theta_1/\theta_o). \qquad (1.80)$$

(iii) The normal distribution. We consider the mean and suppose variance σ^2 to be known. Then we have $E_\theta t(X_1) = E_\theta X_1 = \theta$, $c(\theta) = \theta^2/2\sigma^2$ and $d(\theta) = \theta/\sigma^2$ and we obtain by (1.78)

$$\theta^* = (\theta_o + \theta_1)/2. \qquad (1.81)$$

(iv) The exponential distribution. We suppose $f_\theta(x) = \theta\exp(-\theta x)$, $x \in (0, \infty)$, $\theta \in (0, \infty)$. Then we have $E_\theta t(X_1) = E_\theta X_1 = 1/\theta$, $c(\theta) = -\ln\theta$ and $d(\theta) = -\theta$. Then, by $E_\theta t(X_1) = c'(\theta)/d'(\theta)$ and (1.58) we obtain

$$\frac{1}{\theta^*} = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)} .$$

This implies for $0 < \theta_o, \theta_1 < \infty$, $\theta_o \neq \theta_1$,

$$\theta^* = (\theta_1 - \theta_o)/\ln(\theta_1/\theta_o). \qquad (1.82)$$

We refer to (1.80). ■

Some further consequences of Lemma 1.6.4 and Lemma 1.6.5:

(iii) If Lemma 1.6.5 holds, this lemma can also be used to obtain to given parameters $\theta', \theta_o, \theta_1 \in (\underline{\theta}, \overline{\theta})$, $\theta_o < \theta_1$, a parameter θ'' and a value h so that $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$. In doing this, we have firstly to determine the separating-parameter θ^* which satisfies by Lemma 1.6.4 (ii) the equation

$$\frac{c'(\theta^*)}{d'(\theta^*)} = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)} .$$

If $\theta' \neq \theta^*$ we obtain $\zeta(\theta', \theta^*) > 0$. Then, according to Lemma 1.6.4 (i) there exists a $\theta'' > \theta^*$ which can be obtained as a solution to the equation

$$\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*).$$

As a rule, this equation can be solved by the method of iteration. Finally, the corresponding h can be obtained according to (1.65) by

$$h = \frac{d(\theta'') - d(\theta')}{d(\theta_1) - d(\theta_o)} .$$

Examples, where we may use this approach are, for instance, the distributions considered in Example 1.6.1.

(iv) In many papers on sequential analysis the moment-generating function $\varphi_{Z_{1,\theta_o,\theta_1}}(t)$ of the random variable $Z_{1,\theta_o,\theta_1} = \ln L_{1,\theta_o,\theta_1}$, defined for every $\theta \in \Theta$ by

$$\varphi_{Z_{1,\theta_o,\theta_1}}(t) = E_\theta \exp(tZ_{1,\theta_o,\theta_1}) \text{ for } -\infty < t < +\infty ,$$

plays an important role. For instance, this function can be used to obtain the so-called WALD approximation for the power function and the ASN-function of WALD's likelihood ratio test for the hypothesis H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ based on a sequence $\{X_n\}_{n \in \Gamma^+}$ of i.i.d. random variables. In this context, we refer to Sections 2.1 and 2.7. Especially, a possible non-zero solution t = h to the equation

$$\varphi_{Z_{1,\theta_o,\theta_1}}(t) = 1$$

is then of particular interest. Between the solvability of this equation and conjugacy the following relation holds.

<u>Lemma 1.6.6</u>. Let $\{X_n\}_{n \in \Gamma^+}$ be a sequence of i.i.d. random variables. Suppose that Lemma 1.6.3 holds. Then we have:

(i) If to given $\theta',\theta'',\theta_o,\theta_1 \in (\underline{\theta},\overline{\theta})$ $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ then $t = h$ is a non-zero solution to the equation

$$E_{\theta'}\exp(tZ_{1,\theta_o,\theta_1}) = 1. \tag{1.83}$$

(ii) If to given $\theta',\theta_o,\theta_1 \in (\underline{\theta},\overline{\theta})$, the equation (1.83) has a non-zero solution $t = h$, and if a parameter $\theta'' \in (\underline{\theta},\overline{\theta})$ exists, where

$$d(\theta'') = h(d(\theta_1) - d(\theta_o)) + d(\theta'), \tag{1.84}$$

then we have $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$.

P r o o f. (i) $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ implies $L_{1,\theta',\theta''} = L^h_{1,\theta_o,\theta_1}$ with $h \neq 0$. Hence, we have

$$1 = E_{\theta'}L_{1,\theta',\theta''} = E_{\theta'}L^h_{1,\theta_o,\theta_1} = E_{\theta'}\exp(hZ_{1,\theta_o,\theta_1})$$

and $t = h$ is a non-zero solution to (1.83).

(ii) If $t = h$ is a non-zero solution to (1.83) we obtain

$$E_{\theta'}(f_{\theta_1}(X_1)/f_{\theta_o}(X_1))^h = 1.$$

Since $\theta'' \in (\underline{\theta},\overline{\theta})$ a density $f_{\theta''}(x)$ exists with

$$f_{\theta''}(x) = h(x)\exp(d(\theta'')t(x) - c(\theta'')),\ x \in \mathfrak{X},$$

and we have

$$\int_{\mathfrak{X}} f_{\theta''}(x)d\mu(x) = \int_{\mathfrak{X}} (f_{\theta_1}(x)/f_{\theta_o}(x))^h f_{\theta'}(x)d\mu(x).$$

This, together with (1.84), provides

$$c(\theta'') = h(c(\theta_1) - c(\theta_o)) + c(\theta'). \tag{1.85}$$

Then by (1.84), (1.85) and Lemma 1.6.3 we obtain $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$.∎

2. Likelihood ratio tests

A test (N, δ) for hypothesis H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on the sequence $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ of likelihood ratios is denoted as a likelihood ratio test (LRT). The considerations of Section 1 concerning the sufficiency of the likelihood ratio sequence and its convergence properties for $n \longrightarrow \infty$ already emphasize the importance of the LRTs. For this reason, the following investigations are mainly directed to these tests. We shall see that many properties of tests as they are known for fixed sample size LRTs, see e.g. LEHMANN [53], can be extended to sequential LRTs.

One of the most important properties of a LRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on $\{L_{n,\theta_0,\theta_1}\}_{n\in\Gamma^+}$ is the fact that, cf. Corollary 1.3.1, we can restrict our attention to that class of terminal decision rules depending for every $n \in \Gamma^+$ on the set $\{N = n\}$ only on L_{n,θ_0,θ_1}. Within this class of tests, above all such tests will play an essential role, where also the sample size can be represented only by means of the sequence of likelihood ratios. For a further classification of these tests we introduce the following notations.

To any given sequences of real numbers $\{B_n\}_{n\in\Gamma^+}$ and $\{A_n\}_{n\in\Gamma^+}$, $0 \leq B_n \leq A_n \leq \infty$ for $n \in \Gamma^+$, let N and δ be defined by

$$N = \begin{cases} \inf\{n \geq 1: L_{n,\theta_0,\theta_1} \notin (B_n, A_n)\}, & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise} \end{cases} \tag{2.1}$$

and

$$\delta = \chi_{\{L_{N,\theta_0,\theta_1} \geq A_N,\ N < \infty\}}. \tag{2.2}$$

We denote such an LRT by $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$.

A special version of this LRT is WALD's LRT (WLRT). Then, to given stopping bounds B and A, $0 < B < A < \infty$, the sample size N and the terminal decision rule δ are defined by

$$N = \begin{cases} \inf\{n \geq 1: L_{n,\theta_0,\theta_1} \notin (B, A)\}, & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise} \end{cases} \tag{2.3}$$

and

$$\delta = \chi_{\{L_{N,\theta_0,\theta_1} \geq A,\ N < \infty\}}. \tag{2.4}$$

The properties of this test were systematically investigated by

A. WALD and his co-workers for the first time, and the correspon-
ding results are collected in WALD's famous monograph [77]. Pro-
perties of tests of type $(N,\delta) = \{L_{n,\theta_o,\theta}, B_n, A_n\}_{n\in\Gamma^+}$ were con-
sidered by WEISS [80] and KIEFER, WEISS [51] for the first time.

Sometimes it will be more convenient to consider the logarithm of the likelihood ratio L_{n,θ_o,θ_1}. Let Z_{n,θ_o,θ_1} be defined by

$$Z_{n,\theta_o,\theta_1} = \ln L_{n,\theta_o,\theta_1}, \qquad n\in\Gamma^+,$$

and let b_n and a_n be defined by

$$b_n = \ln B_n \quad \text{and} \quad a_n = \ln A_n, \qquad n\in\Gamma^+,$$

respectively. Then we obtain instead of (2.1) and (2.2)

$$N = \begin{cases} \inf\{n \geq 1 : Z_{n,\theta_o,\theta_1} \notin (b_n, a_n)\}, & \text{if such an n exists,} \\ \infty & \text{, otherwise} \end{cases}$$

and

$$\delta = \chi_{\{Z_{N,\theta_o,\theta_1} \geq a_N,\ N<\infty\}}.$$

Then we shall also write $(N,\delta) = \{Z_{n,\theta_o,\theta_1}, b_n, a_n\}_{n\in\Gamma^+}$.

<u>Example 2.1.0</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a WLRT for $H_o: \theta = \theta_o$ against $H_1: \theta = \theta_1$, $\theta_o \neq \theta_1$, based on a sequence of i.i.d. random variables $\{X_n\}_{n\in\Gamma^+}$ having density

$$f_\theta(x) = h(x)\exp(d(\theta)t(x) - c(\theta)),\ x\in\mathfrak{X},\ \theta\in\Theta.$$

Then we obtain

$$L_{n,\theta_o,\theta_1} = \exp((d(\theta_1) - d(\theta_o))\sum_{i=1}^{n} t(X_i) - (c(\theta_1) - c(\theta_o))n)$$

and

$$Z_{n,\theta_o,\theta_1} = (d(\theta_1) - d(\theta_o))\sum_{i=1}^{n} t(X_i) - (c(\theta_1) - c(\theta_o))n, \quad n\in\Gamma^+$$

If b and a are defined by

$$b = \ln B \quad \text{and} \quad a = \ln A,$$

respectively, then we can describe the test $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ as follows. We continue sampling as long as for $n = 1,2,\dots$ the variables Z_{n,θ_o,θ_1} satisfy the so-called <u>critical inequalities</u>

$$b < Z_{n,\theta_o,\theta_1} < a. \tag{2.5}$$

We stop sampling at stage n if

$$Z_{n,\theta_o,\theta_1} \leq b \quad \text{or} \quad Z_{n,\theta_o,\theta_1} \geq a$$

holds at this stage for the first time. If $Z_{n,\theta_o,\theta_1} \leq b$, then we accept hypothesis H_o, in case of $Z_{n,\theta_o,\theta_1} \geq a$ we reject hypothesis H_o or accept H_1 (see Fig. 2.1), respectively.

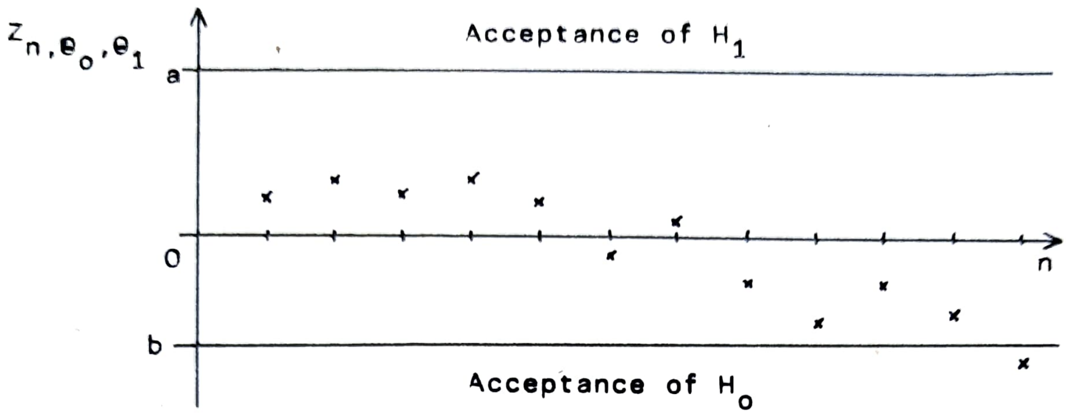

Fig. 2.1 Graphical representation of a WLRT

If $\theta_o < \theta_1$ and if d is strictly monotonically increasing in θ on Θ the critical inequalities (2.5) reduce to

$$sn + h_a < \sum_{i=1}^{n} t(X_i) < sn + h_r, \quad n \in \Gamma^+, \tag{2.6}$$

where

$$h_a = \frac{b}{d(\theta_1) - d(\theta_o)} \quad \text{and} \quad h_r = \frac{a}{d(\theta_1) - d(\theta_o)}, \tag{2.7}$$

and

$$s = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)}. \tag{2.8}$$

The corresponding graph is shown in Fig. 2.2.

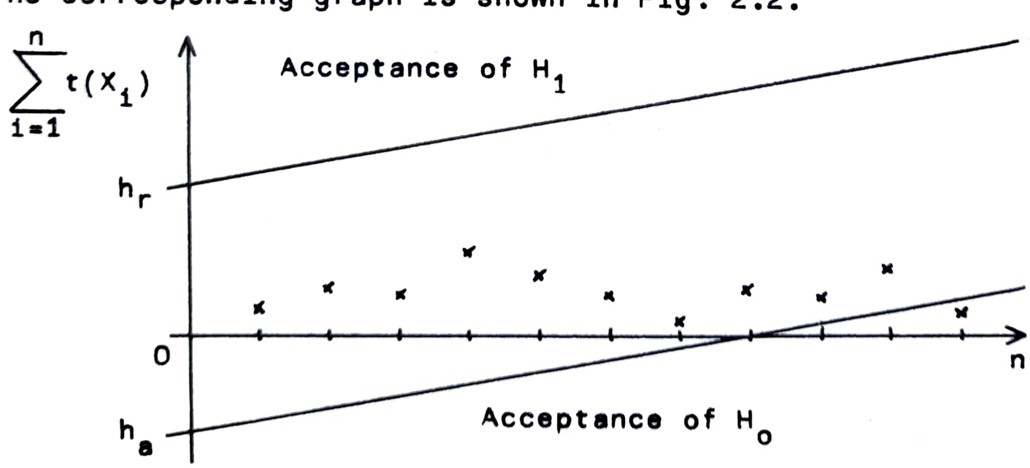

Fig. 2.2 Further graphical representation of a WLRT

In practice, each sample point $(n, \sum_{i=1}^{n} t(x_i))$ for $n = 1,2,\ldots$ can be plotted in this plane and we continue sampling as long as these sample points are contained in the continue-sampling region. This region lies between two lines having equal slopes and the intercepts h_a and h_r.

We notice, if Lemma 1.6.4 is applicable, then we have

$$Z_{n,\theta_o,\theta_1} = (d(\theta_1) - d(\theta_o))\left(\sum_{i=1}^{n} t(X_i) - n\frac{c'(\theta^*)}{d'(\theta^*)}\right), \quad n\in\Gamma^+,$$

where the so-called separating-parameter θ^* is determined by (1.58). Hence, by (1.58) and (2.8) we obtain

$$s = c'(\theta^*)/d'(\theta^*).$$

If, moreover, $E_\theta t(X_1) = \theta$ holds, then we have by (1.60) even

$$s = \theta^*$$

so that for this case the slope of the acceptance line and the rejection line, respectively, considered in Fig. 2.2, is equal to the separating-parameter θ^*.

Some particular examples (see also Example 1.6.1):

(i) The binomial proportion. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. Bernoulli variables with

$$f_\theta(x) = \theta^x(1-\theta)^{1-x}, \quad x\in\{0,1\}, \quad \theta\in(0,1).$$

Consider WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$, $0<\theta_o<\theta_1<1$. Then, for every $n\in\Gamma^+$, we have

$$L_{n,\theta_o,\theta_1} = \left(\frac{\theta_1}{\theta_o}\right)^{\sum_{i=1}^{n} X_i}\left(\frac{1-\theta_1}{1-\theta_o}\right)^{n-\sum_{i=1}^{n} X_i}$$

and

$$Z_{n,\theta_o,\theta_1} = \left(\ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}\right)\sum_{i=1}^{n} X_i + \left(\ln\frac{1-\theta_1}{1-\theta_o}\right)n,$$

and critical inequality (2.5) reduces to (2.6) with

$$h_a = \ln B\Big/\ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}, \qquad h_r = \ln A\Big/\ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}$$

and

$$s = \ln\frac{1-\theta_o}{1-\theta_1}\Big/\ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}.$$

If $\theta_o = \frac{1}{2} - \varepsilon$ and $\theta_1 = \frac{1}{2} + \varepsilon$ with $\varepsilon \in (0, \frac{1}{2})$, then we obtain

$$Z_{n,\theta_o,\theta_1} = 2(\ln \xi_\varepsilon) \sum_{i=1}^{n} X_i - (\ln \xi_\varepsilon)\, n,$$

where ξ_ε is defined by

$$\xi_\varepsilon = (1+2\varepsilon)/(1-2\varepsilon).$$

Further, we obtain

$$h_a = \ln B/2 \ln \xi_\varepsilon, \quad h_r = \ln A/2 \ln \xi_\varepsilon \quad \text{and} \quad s = 1/2.$$

(ii) The Poisson mean. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables with density

$$f_\theta(x) = (\theta^x/x!)\exp(-\theta), \quad x \in \Gamma_o^+, \quad \theta \in (0, \infty).$$

Consider WLRT $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$, $0 < \theta_o < \theta_1 < \infty$. Then

$$L_{n,\theta_o,\theta_1} = (\theta_1/\theta_o)^{\sum_{i=1}^{n} X_i} \exp(-n(\theta_1 - \theta_o))$$

and

$$Z_{n,\theta_o,\theta_1} = \left(\ln(\theta_1/\theta_o)\right)\sum_{i=1}^{n} X_i - n(\theta_1 - \theta_o),$$

$n \in \Gamma^+$, and critical inequality (2.5) reduces to (2.6) with

$$h_a = \ln B/\ln(\theta_1/\theta_o), \quad h_r = \ln A/\ln(\theta_1/\theta_o) \quad \text{and}$$

$$s = (\theta_1 - \theta_o)/\ln(\theta_1/\theta_o).$$

(iii) The normal mean. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables with

$$f_\theta(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right), \quad \theta, x \in (-\infty, +\infty),$$

where σ^2 is known. Consider WLRT $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$, $-\infty < \theta_o < \theta_1 < +\infty$. Then

$$L_{n,\theta_o,\theta_1} = \exp\left(\frac{\theta_1 - \theta_o}{\sigma^2}\sum_{i=1}^{n} X_i - \frac{\theta_1^2 - \theta_o^2}{2\sigma^2}\right)$$

and

$$Z_{n,\theta_o,\theta_1} = \frac{\theta_1 - \theta_o}{\sigma^2} \sum_{i=1}^{n} X_i - \frac{\theta_1^2 - \theta_o^2}{2\sigma^2} \cdot n,$$

$n \in \Gamma^+$, and critical inequality (2.5) reduces to (2.6) with

$$h_a = (\sigma^2 \ln B)/(\theta_1 - \theta_o), \quad h_r = (\sigma^2 \ln A)/(\theta_1 - \theta_o)$$

and

$$s = (\theta_0 + \theta_1)/2.$$

(iv) <u>The exponential mean</u>. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables with

$$f_\theta(x) = \theta \exp(-\theta x), \quad x \in (0,\infty), \ \theta \in (0,\infty).$$

Consider WLRT $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$, $0 < \theta_0 < \theta_1 < \infty$. Then

$$L_{n,\theta_0,\theta_1} = (\theta_1/\theta_0)^n \exp\left(-(\theta_1 - \theta_0)\sum_{i=1}^{n} X_i\right), \quad n\in\Gamma^+$$

$$Z_{n,\theta_0,\theta_1} = -(\theta_1 - \theta_0)\sum_{i=1}^{n} X_i + n\ln(\theta_1/\theta_0),$$

$n\in\Gamma^+$. and critical inequality (2.5) reduces to

$$sn + h_a > \sum_{i=1}^{n} X_i > sn + h_r \tag{2.9}$$

with

$$h_a = -(\ln B)/(\theta_1 - \theta_0), \quad h_r = -(\ln A)/(\theta_1 - \theta_0)$$

and

$$s = (\ln(\theta_1/\theta_0))/(\theta_1 - \theta_0).$$

We note that here H_0 is accepted or rejected as the lower or the upper inequality in (2.9) is violated for the first time.■

The WLRT possesses a comparatively simple structure. Nevertheless, it possesses rather far-reaching optimality properties. In this context we refer to Sections 2.2 and 2.8.

In view of the computation ot the characteristics of a WLRT, the so-called fundamental identity or WALD's identity used to be the main device in investigating the properties of a WLRT (see e.g.[77]). By means of this identity Wald obtained his famous approximations for the power function, the OC-function, the ASN-function and the stopping bounds of a WLRT. Here we do not follow this classical way but use the concept of conjugacy introduced in Section 1.6. by which more general assertions on LRTs and, especially WLRTs, can be obtained.

2.1 The power function

The power function and its counterpart, the OC-function, are the most important characteristics for assessing the statistical properties of the terminal decision rule of a test. They provide

information about the probability of the acceptance of hypothesis H_1 or H_0, respectively, at a finite sampling stage, depending on the parameter $\theta \in \Theta$. The power function $M(\theta)$ and the OC-function $Q(\theta)$ of a given test (N, δ) are defined by

$$M(\theta) = E_\theta \delta \chi_{\{N < \infty\}}, \quad \theta \in \Theta, \tag{2.10}$$

and

$$Q(\theta) = E_\theta (1 - \delta) \chi_{\{N < \infty\}}, \quad \theta \in \Theta. \tag{2.11}$$

According to this definition, we obtain

$$M(\theta) + Q(\theta) \leq 1, \quad \theta \in \Theta,$$

where the strict equality holds for every $\theta \in \Theta$ iff (N, δ) is closed. If (N, δ) is a closed test, then it is sufficient to consider only one of these characteristics.

It is a very difficult problem to obtain general assertions on the power function and the OC-function of a given test (N, δ) without any structural assumptions on sample size N and terminal decision rule δ. Likewise the possibilities of the computation of certain characteristics of a test (N, δ) essentially depend on the structure of N and δ. We start our investigations with a theorem which provides relations between certain conditional expectation values and the power function and the OC-function for an arbitrary test (N, δ).

Theorem 2.1.1. Let (N, δ) be any given LRT for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. Denote by F_0 and F_1 the events of the acceptance of H_0 and H_1, respectively, where $F_0, F_1 \in \mathfrak{F}_N$ and $F_0, F_1 \subseteq \{N < \infty\}$. If $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$, then

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1} \chi_{\{N < \infty\}} \mid F_0) = Q(\theta'')/Q(\theta') \tag{2.12}$$

and

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1} \chi_{\{N < \infty\}} \mid F_1) = M(\theta'')/M(\theta'). \tag{2.13}$$

Proof. With $\varphi = 1$ by Lemma 1.6.1 we obtain

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1} \chi_{\{N<\infty\}} \mid F_0) = (P_{\theta''}(F_0)/P_{\theta'}(F_0)) E_{\theta''}(\chi_{\{N<\infty\}} \mid F_0). \tag{2.14}$$

Then, by definition of the OC-function, we obtain

$$P_\theta(F_0) = E_\theta(1 - \delta)\chi_{\{N < \infty\}} = Q(\theta), \quad \theta \in \Theta.$$

Furthermore, since $F_0 \subseteq \{N < \infty\}$ we have

$$E_{\theta''}(\chi_{\{N<\infty\}} \mid F_0) = 1.$$

This together with (2.14) provides (2.12). Analogously we obtain (2.13). ■

By means of this theorem we obtain the following general formula for the power function of an LRT. Denote by $B(\theta',\theta'')$ and $A(\theta',\theta'')$ the left-hand sides of (2.12) and (2.13), respectively.

Corollary 2.1.1. Suppose the assumptions of Theorem 2.1.1 are fulfilled. If N is closed, then

$$M(\theta') = \frac{1 - B(\theta',\theta'')}{A(\theta',\theta'') - B(\theta',\theta'')} \tag{2.15}$$

and

$$M(\theta'') = A(\theta',\theta'')M(\theta'). \tag{2.16}$$

Proof. If N is closed, we have $M(\theta) + Q(\theta) = 1$, $\theta \in \Theta$. Hence, instead of (2.12), we obtain

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1}\chi_{\{N<\infty\}}|F_0) = (1 - M(\theta''))/(1 - M(\theta')).$$

This, together with (2.13), implies (2.15) and (2.16). ■

Thus, in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ the computation of $M(\theta')$ and $M(\theta'')$ of any given closed LRT (N,δ) for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ can be reduced to the computation of the conditional expectation values

$$B(\theta',\theta'') = E_{\theta'}(L^h_{N,\theta_0,\theta_1}|F_0)$$

and

$$A(\theta',\theta'') = E_{\theta'}(L^h_{N,\theta_0,\theta_1}|F_0).$$

2.1.1 The WALD approximation for the power function of a WLRT

Applying Corollary 2.1.1 to WLRTs and taking advantage of the concept of conjugacy we obtain a formula for the power function of a WLRT as follows.

Lemma 2.1.0. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ be a closed WLRT, $0<B<1<A<\infty$. If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$, where

$$P_{\theta'}(L_{N,\theta_0,\theta_1} = B|F_0) = 1 \tag{2.17}$$

and

$$P_{\theta'}(L_{N,\theta_0,\theta_1} = A|F_1) = 1, \tag{2.18}$$

then

$$M(\theta') = (1 - B^h)/(A^h - B^h) \tag{2.19}$$

and

$$M(\theta'') = A^h(1 - B^h)/(A^h - B^h). \tag{2.20}$$

Proof. Because of (2.17) and (2.18) we obtain

$$B(\theta',\theta") = B^h \quad \text{and} \quad A(\theta',\theta") = A^h.$$

Then, (2.19) and (2.20) follows immediately from (2.15) and (2.16).■

Let $M^*(h)$ be a real function defined to given stopping bounds B and A, $0<B<1<A<\infty$, by

$$M^*(h) = \begin{cases} (1 - B^h)/(A^h - B^h) & \text{for } -\infty<h<+\infty \text{ and } h \neq 0, \\ (-\ln B)/(\ln A - \ln B) & \text{for } h = 0. \end{cases} \quad (2.21)$$

Then $M^*(h)$ is a continuous function in h on $(-\infty,+\infty)$, and we obtain in case of $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ instead of (2.19) and (2,20) also

$$M(\theta') = M^*(h) \quad (2.22)$$

and

$$M(\theta") = M^*(-h), \quad (2.23)$$

respectively.

Of course, the assumptions (2.17) and (2.18) are quite restrictive. If, instead of (2.17) and (2.18), we only have

$$P_{\theta'}(B - \varepsilon \leq L_{N,\theta_0,\theta_1} \leq B | F_0) = 1 \quad (2.24)$$

and

$$P_{\theta"}(A \leq L_{N,\theta_0,\theta_1} \leq A + \varepsilon \mid F_1) = 1 \quad (2.25)$$

for any given sufficiently small $\varepsilon > 0$ then, instead of (2.22), we obtain

$$M(\theta') \approx M^*(h) \quad \text{if} \quad (\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1). \quad (2.26)$$

This is the so-called WALD approximation of the power function of a WLRT. We shall call $B - L_{N,\theta_0,\theta_1}$ and $L_{N,\theta_0,\theta_1} - A$ the excess of L_{N,θ_0,θ_1} at termination over B and A, respectively. If (2.24) and (2.25) hold for a small $\varepsilon > 0$, we shall say the excess of L_{N,θ_0,θ_1} over B and A is small for the given $\theta \in \Theta$.

Example 2.1.1. Continuation of Example 2.1.0. In order to obtain the WALD approximations of the power functions of the tests considered there, we need the corresponding relations between θ' and h in case of $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$. According to Lemma 1.6.3, we have $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ iff

$$d(\theta") - d(\theta') = h(d(\theta_1) - d(\theta_0))$$

and

$$c(\theta") - c(\theta') = h(c(\theta_1) - c(\theta_0)).$$

For a given θ' this system of equations can be used to determine

the required parameter $h \neq 0$ if such an h exists. As a rule, we obtain an iteration formula for h. Vice versa, if $h \neq 0$ is given, as a rule, the above systems of equations provides an explicit formula for the corresponding θ'. The details have already been investigated in Example 1.6.1.

An example where the excess of L_{N,θ_0,θ_1} at termination over B and A is zero at least for some values of the stopping bounds B and A is the following.

<u>Example 2.1.2</u>. Let $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a WLRT for

$$H_0: \theta = \theta_0 = \tfrac{1}{2} - \varepsilon \text{ against } H_1: \theta = \theta_1 = \tfrac{1}{2} + \varepsilon,\ 0<\varepsilon<\tfrac{1}{2},$$

based on a sequence of independent Bernoulli distributed random variables (cf. Example 2.1.0 (i)). Then we have

$$Z_{n,\theta_0,\theta_1} = 2\left(\ln \frac{1+2\varepsilon}{1-2\varepsilon}\right)\sum_{i=1}^{n} X_i - \left(\ln \frac{1+2\varepsilon}{1-2\varepsilon}\right) n,\ n\in\Gamma^+.$$

so that Z_{n,θ_0,θ_1} is an integer multiple of $\xi_\varepsilon = \ln((1+2\varepsilon)/(1-2\varepsilon))$ for every $n\in\Gamma^+$. Hence, relations (2.17) and (2.18) hold iff integers k_a and k_r, $k_a < 0 < k_r$, exist, where

$$\ln B = k_a \xi_\varepsilon \quad \text{and} \quad \ln A = k_r \xi_\varepsilon . \tag{2.27}$$

The power function: According to Example 1.6.1 (i) (cf. (1.41) and (1.40)) we obtain $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ if

$$h = \ln((1-\theta')/\theta')/\xi_\varepsilon$$

and $\theta'' = 1 - \theta'$. Thus, by (2.22), we obtain

$$M(\theta') = M^*(h) = \frac{1 - \left(\frac{1-\theta'}{\theta'}\right)^{k_a}}{\left(\frac{1-\theta'}{\theta'}\right)^{k_r} - \left(\frac{1-\theta'}{\theta'}\right)^{k_a}} \quad \text{for } \theta \neq \tfrac{1}{2} \tag{2.28}$$

in case of (2.27).■

The case is still open where to any given $\theta' \in \Theta$ a $\theta'' \neq \theta'$ does not exist so that $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$. We shall say then θ' is an <u>exceptional point</u>. In order to consider this situation we suppose that the following continuity assumptions are fulfilled.

(i) $M(\theta)$ is a continuous function of $\theta \in \Theta$.

(ii) For every sequence $\{\theta_n'\}_{n\in\Gamma^+} \in \Theta$ with $\lim_{n\to\infty} \theta_n' = \theta'$ there exists

a sequence $\{\theta''_n\}_{n\in\Gamma^+}\in\Theta$ and a sequence $\{h_n\}_{n\in\Gamma^+}\in R^1$ with $\lim_{n\to\infty} h_n = h$, where $(\theta'_n,\theta''_n) \overset{h_n}{\sim} (\theta_o,\theta_1)$, $h_n \neq 0$, $n\in\Gamma^+$.

Then, if the excess at termination over B and A, respectively, is zero, by (2.21) we obtain

$$M(\theta') = \lim_{\theta'_n\to\theta'} M(\theta'_n) = \lim_{h_n\to h} M^*(h_n) = M^*(h).$$

If now for any given sequence $\{\theta'_n\}_{n\in\Gamma^+}\in\Theta$ with $\lim_{n\to\infty}\theta'_n = \theta^*$ we have $\lim_{n\to\infty} h_n = 0$, and if according to (2.21) for $h = 0$ $M^*(h)$ is defined by

$$M^*(0) = \lim_{h\to 0} M^*(h) = (-\ln B)/(\ln A - \ln B)$$

then, under the above continuity assumptions, we obtain

$$M(\theta^*) = \lim_{\theta'_n\to\theta^*} M(\theta'_n) = \lim_{h_n\to 0} M^*(h_n) = M^*(0)$$

$$= (-\ln B)/(\ln A - \ln B).$$

Hence, if the excess at termination over B and A, respectively, is zero, we obtain the following parametric form for the power function of a WLRT under the above continuity assumptions:

$$M(\theta') = \begin{cases} (1 - B^h)/(A^h - B^h), & \text{if } (\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1), \\ (-\ln B)/(\ln A - \ln B), & \text{if } \theta' \text{ is an exceptional point.} \end{cases} \tag{2.29}$$

We remark that the above continuity assumptions are fulfilled, for instance, for the exponential family considered in Lemma 1.6.4. For this family, the parametric form of the power function can be modified as follows.

<u>C o r o l l a r y 2.1.2.</u> Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT based on $\{X_n\}_{n\in\Gamma^+}$ where Lemma 1.6.4 holds. Suppose the excess of L_{N,θ_o,θ_1} at termination is zero. Then, for every $\theta'\in(\underline{\theta},\bar{\theta})$, we have

$$M(\theta') = M^*(h) \quad \text{with} \quad h = \frac{d(\theta'') - d(\theta')}{d(\theta_1) - d(\theta_o)}, \tag{2.30}$$

where parameter θ'' is given by

$$\zeta(\theta',\theta^*) = \zeta(\theta'',\theta^*)$$

and θ^* denotes the separating-parameter given by

$$\frac{c'(\theta^*)}{d'(\theta^*)} = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)} .$$

P r o o f. This immediately follows from Lemma 2.1.1, Lemma 1.6.5 and the definition of $M^*(h)$ according to (2.21).■

If we may only assume that the excess of L_{N,θ_o,θ_1} over B and A at termination is small, then (2.30) holds approximately and the right-hand side of (2.30) is again the so-called WALD approximation for the power function $M(\theta)$.

2.1.2 The WALD approximations for the stopping bounds of a WLRT

Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT. If (2.17) and (2.18) holds for $\theta' = \theta_o$ then (2.19) and (2.20) imply

$$M(\theta_o) = \frac{1 - B}{A - B} \quad \text{and} \quad M(\theta_1) = A\,\frac{1 - B}{A - B},$$

and to given α and β, $0 < \alpha, \beta < 1$, we can always choose A and B in such a manner that

$$\frac{1 - B}{A - B} = \alpha \quad \text{and} \quad A\,\frac{1 - B}{A - B} = 1 - \beta$$

holds. This implies

$$B = B^* = \frac{\beta}{1 - \alpha} \quad \text{and} \quad A = A^* = \frac{1 - \beta}{\alpha} \tag{2.31}$$

with $0 < B^* < 1 < A^* < \infty$ for $\alpha + \beta < 1$. If then $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B^*, A^*\}_{n\in\Gamma^+}$ is closed and if

$$P_{\theta_o}(L_{N,\theta_o,\theta_1} = B^* \mid F_o) = 1 \text{ and } P_{\theta_o}(L_{N,\theta_o,\theta_1} = A^* \mid F_1) = 1 \tag{2.32}$$

holds, the power function of this test satisfies

$$M(\theta_o) = \alpha \quad \text{and} \quad M(\theta_1) = 1 - \beta. \tag{2.33}$$

If the excess of L_{N,θ_o,θ_1} over B and A is small, we obtain instead of (2.33)

$$M(\theta_o) \approx \alpha \quad \text{and} \quad M(\theta_1) \approx 1 - \beta$$

and the stopping bounds B^* and A^*, given by (2.31), are denoted as the <u>WALD approximations</u> for the stopping bounds of a WLRT.

2.1.3 A test for hypothesis $H_0: \theta \leq \theta^*$ against $H_1: \theta > \theta^*$

Let $\{X_n\}_{n \in \Gamma^+}$ be a sequence of i.i.d. random variables having density

$$f_\theta(x) = h(x)\exp(d(\theta)t(x) - c(\theta)), \quad x \in \mathfrak{X}, \quad \theta \in (\underline{\theta}, \overline{\theta}). \quad (2.34)$$

Suppose that Lemma 1.6.4 holds. Our aim is to discriminate between hypotheses

$$H_0: \theta \leq \theta^* \quad \text{and} \quad H_1: \theta > \theta^*, \quad \underline{\theta} < \theta^* < \overline{\theta}. \quad (2.35)$$

In doing this, we consider a test $(\hat{N}, \hat{\delta})$ defined as follows. To given stopping bounds $\hat{B}$ and $\hat{A}$, $0 < \hat{B} < 1 < \hat{A} < \infty$, let the sample size $\hat{N}$ and the terminal decision rule $\hat{\delta}$ be defined by

$$\hat{N} = \begin{cases} \inf\{n \geq 1: L_n \notin (B, \hat{A})\}, & \text{if such an } n \text{ exists}, \\ \infty & \text{, otherwise} \end{cases} \quad (2.36)$$

and

$$\hat{\delta} = \chi_{\{L_N^* \geq \hat{A},\ \hat{N} < \infty\}}, \quad (2.37)$$

where L_n^* is given by

$$L_n^* = \exp\left(\sum_{i=1}^{n} t(X_i) - n\,\frac{c'(\theta^*)}{d'(\theta^*)}\right), \quad n \in \Gamma^+. \quad (2.38)$$

The heuristical background for choosing $\hat{N}$ and $\hat{\delta}$ in such a manner is the following. As already stated above, for family (2.34) we have

$$E_\theta t(X_1) = c'(\theta)/d'(\theta), \quad \theta \in (\underline{\theta}, \overline{\theta}),$$

where $c'(\theta)/d'(\theta)$ is strictly monotonically increasing in θ on $(\underline{\theta}, \overline{\theta})$. Then, for any given $\theta^* \in (\underline{\theta}, \overline{\theta})$, we obtain

$$E_\theta t(X_1) - E_{\theta^*} t(X_1) < 0 \quad \text{for} \quad \theta < \theta^*$$

and

$$E_\theta t(X_1) - E_{\theta^*} t(X_1) > 0 \quad \text{for} \quad \theta > \theta^*.$$

That means, if for sufficiently large values of n

$$\sum_{i=1}^{n} t(X_i) - n\,E_{\theta^*} t(X_1) < 0,$$

then this is a possible hint to the fact that the true parameter θ satisfies $\theta < \theta^*$. Conversely, positive values of $\sum_{i=1}^{n} t(X_i) - nE_{\theta^*} t(X_1)$ may be an indication of $\theta^* < \theta$. This motivates the choice of $\hat{N}$ and $\hat{\delta}$ according to (2.36).

Test $(\hat{N}, \hat{\delta})$ is a WLRT in the following sense.

L e m m a 2.1.1. Let $(\hat{N}, \hat{\delta})$ be a test for H_0: $\theta \leq \theta^*$ against H_1: $\theta > \theta^*$, $\underline{\theta} < \theta^* < \overline{\theta}$, defined by (2.36), (2.37) and (2.38). We suppose that Lemma 1.6.4 holds. Then, for every pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$, $\theta' \neq \theta''$, satisfying

$$\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*) > 0 \qquad (2.39)$$

test $(\hat{N}, \hat{\delta})$ is identical with test

$$(N, \delta) = \{L_{n,\theta',\theta''}, B^{d_{\theta',\theta''}}, A^{d_{\theta',\theta''}}\}_{n \in \Gamma^+}, \qquad (2.40)$$

where

$$d_{\theta',\theta''} = d(\theta'') - d(\theta'). \qquad (2.41)$$

P r o o f. According to Lemma 1.6.4, a continuum of pairs $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$ exists for every $\theta^* \in (\underline{\theta}, \overline{\theta})$ so that (2.39) holds with $\theta' < \theta^* < \theta''$. Hence, a corresponding $\theta'' > \theta^*$ will exist for every $\theta' < \theta^*$ so that (2.39) holds. For every pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$ satisfying (2.39) we have

$$\frac{c(\theta'') - c(\theta')}{d(\theta'') - d(\theta')} = \frac{c'(\theta^*)}{d'(\theta^*)}.$$

Thus we obtain

$$L_{n,\theta',\theta''} = \exp((d(\theta'') - d(\theta'))\sum_{i=1}^{n} X_i - n(c(\theta'') - c(\theta'))$$

$$= \exp((d(\theta'') - d(\theta'))\left(\sum_{i=1}^{n} X_i - n\,\frac{c'(\theta^*)}{d'(\theta^*)}\right)$$

$$= (L_n^*)^{d(\theta'') - d(\theta')}, \quad n \in \Gamma^+.$$

If now $d_{\theta',\theta''}$ is defined by (2.41), then the critical inequality

$$\hat{B} < L_n^* < \hat{A}$$

of test $(\hat{N}, \hat{\delta})$ is equivalent to

$$B^{d_{\theta',\theta''}} < L_{n,\theta',\theta''} < A^{d_{\theta',\theta''}},$$

$n \in \Gamma^+$. Furthermore, inequality $L_n^* \leq B$ is equivalent to $L_{n,\theta',\theta''} \leq B^{d_{\theta',\theta''}}$, $n \in \Gamma^+$. Hence, tests $(\hat{N}, \hat{\delta})$ and (N, δ) are identical.■

An immediate conclusion of this lemma is that test $(\hat{N}, \hat{\delta})$ will have the same optimality properties for $\theta = \theta'$ and $\theta = \theta''$ like WLRT (2.40) Consequently, the computation of the characteristics of test $(\hat{N}, \hat{\delta})$ can be reduced to the computation of the characteristics of WLRT (2.40).

We consider the power function $\hat{M}(\theta)$ of $(\hat{N}, \hat{\delta})$.

L e m m a 2.1.2. Let $(\hat{N}, \hat{\delta})$ be a closed test for H_0: $\theta \le \theta^*$ against H_1: $\theta > \theta^*$, $\underline{\theta} < \theta^* < \bar{\theta}$, defined by (2.36), (2.37) and (2.38). We suppose that Lemma 1.6.4 holds. To any given $\theta' \in (\underline{\theta}, \bar{\theta})$ let θ'' be defined by

$$\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*). \tag{2.42}$$

Further we suppose that

$$P_{\theta'}(L_N^* = \hat{B} | H_0 \text{ is accepted}) = 1 \tag{2.43}$$

and

$$P_{\theta'}(L_N^* = \hat{A} | H_1 \text{ is accepted}) = 1 \tag{2.44}$$

holds. Then we have

$$\hat{M}(\theta') = M^*(d_{\theta',\theta''}) \tag{2.45}$$

with

$$d_{\theta',\theta''} = d(\theta'') - d(\theta'), \tag{2.46}$$

where M^* is defined by (2.21).

P r o o f. By Lemma 2.1.1 test $(\hat{N}, \hat{\delta})$ is identical with test (N, δ), defined by (2.40). If to given $\theta' \neq \theta^*$ the corresponding θ'' is determined by (2.42), we have $\theta'' \neq \theta^*$, and, likewise to the proof of Lemma 2.1.1, we obtain

$$L_{n,\theta',\theta''} = (L_n^*)^{d_{\theta',\theta''}}, \quad n \in \Gamma^*,$$

so that (2.43) and (2.44) are equivalent to

$$P_{\theta'}(L_{N,\theta',\theta''} = \hat{B}^{d_{\theta',\theta''}} | H_0 \text{ is accepted}) = 1 \tag{2.47}$$

and

$$P_{\theta'}(L_{N,\theta',\theta''} = \hat{A}^{d_{\theta',\theta''}} | H_1 \text{ is accepted}) = 1. \tag{2.48}$$

Denote by $M(\theta)$ the power function of (N, δ). Then (2.47), (2.48), $(\theta', \theta'') \overset{1}{\sim} (\theta', \theta'')$, (2.19), (2.21) and Lemma 2.1.1 imply

$$\hat{M}(\theta') = M(\theta') = \frac{1 - B^{d_{\theta',\theta''}}}{A^{d_{\theta',\theta''}} - B^{d_{\theta',\theta''}}} = M^*(d_{\theta',\theta''}) \text{ for } \theta' \neq \theta^*.$$

Further, if θ' tends to θ^* from below, the corresponding θ'' given by (2.42) tends to θ^* from above, and we obtain $\lim_{\theta' \to \theta^*} d_{\theta',\theta''} = d_{\theta^*,\theta^*} = 0$. This implies

$$\hat{M}(\theta^*) = \lim_{\theta' \to \theta^*} M(\theta') = \lim_{d_{\theta',\theta''} \to 0} \frac{1 - B^{d_{\theta',\theta''}}}{A^{d_{\theta',\theta''}} - B^{d_{\theta',\theta''}}}$$

$$= (-\ln B)/(\ln A - \ln B),$$

which completes the proof.■

If instead of (2.43) and (2.44) we may only suppose that the excess of L_N^* at termination over B and A is small, we obtain

$$\hat{M}(\theta') \approx M^*(d_{\theta',\theta''})$$

instead of (2.45), which can be considered as the WALD approximation for the power function of test $(\hat{N},\hat{\delta})$.

We now turn to the choice of the stopping bounds $\hat{B}$ and $\hat{A}$ of test $(\hat{N},\hat{\delta})$. In general, it will be not possible to choose $\hat{B}$ and $\hat{A}$ in such a manner that for an arbitrary pair $\theta_0,\theta_1 \in (\underline{\theta},\overline{\theta})$, $\theta_0 < \theta^* < \theta_1$,

$$\hat{M}(\theta_0) = \alpha \quad \text{and} \quad \hat{M}(\theta_1) = 1 - \beta$$

holds to given α,β, $0 < \alpha,\beta < 1$, $\alpha + \beta < 1$. The reason is that to given $\theta^* \in (\underline{\theta},\overline{\theta})$ (2.42) holds only for selected parameter pairs θ_0,θ_1. Further, a test for testing hypotheses (2.35) should possess the following property. If the true parameter, say θ', satisfies $\theta' < \theta^*$, then the probability of the acceptance of H_0 should be larger than the probability of the acceptance of H_1. Conversely, if $\theta^* < \theta'$, the probability for a decision for H_1 should be larger then the probability for a decision for H_0. Hence, if we may suppose that $(\hat{N},\hat{\delta})$ is closed, a reasonable requirement for our test is

$$\hat{M}(\theta^*) = 1/2.$$

Then the stopping bounds $\hat{B}$ and $\hat{A}$ can be chosen as follows.

L e m m a 2.1.3. Suppose that Lemma 2.1.2 holds. If to a given θ', $\underline{\theta} < \theta' < \theta^*$, and α, $0 < \alpha < 1/2$, the stopping bounds $\hat{B}$ and $\hat{A}$ of $(\hat{N},\hat{\delta})$ are chosen by

$$\hat{B} = \left(\frac{\alpha}{1-\alpha}\right)^{1/d_{\theta',\theta''}} \tag{2.49}$$

and

$$\hat{A} = 1/\hat{B}, \tag{2.50}$$

where θ'' and $d_{\theta',\theta''}$ are determined by (2.42) and (2.46), respectively, then the power function $\hat{M}(\theta)$ of $(\hat{N},\hat{\delta})$ satisfies

$$\hat{M}(\theta^*) = 1/2 \tag{2.51}$$

and

$$\hat{M}(\theta') = \alpha. \tag{2.52}$$

P r o o f. For $\theta' = \theta^*$ by (2.45) we obtain

$$\hat{M}(\theta^*) = (-\ln\hat{B})/(\ln\hat{A} - \ln\hat{B})$$

and we have $\hat{M}(\theta^*) = 1/2$ iff $\hat{A} = 1/\hat{B}$. This and (2.45) imply

$$\hat{M}(\theta') = \hat{B}^{d_{\theta',\theta''}}/(1 + \hat{B}^{d_{\theta',\theta''}})$$

for every $\theta' < \theta^*$. Hence, we have $\hat{M}(\theta') = \alpha$ iff (2.49) holds.■

We note that we obtain

$$\hat{M}(\theta'') = 1 - \alpha \tag{2.53}$$

for the corresponding θ'', given by (2.42).
If, instead of (2.43) and (2.44), we may only suppose that the excess of L_N^* at termination over $\hat{B}$ and $\hat{A}$ is small, the relations (2.51), (2.52) and (2.53) hold only approximately and are again approximations in the sense of the WALD approximations.

We remark that the above approach of the derivation of the power function $\hat{M}(\theta)$ of test $(\hat{N}, \hat{\delta})$ can be considered simultaneously as an alternative approach deriving the WALD approximations for the power function and the stopping bounds of a WLRT for the exponential family (2.34). An advantage of the approach considered here is, that it is not necessary to determine the conjugacy parameter h explicitly.

A further possibility of the choice of the stopping bounds for test $(\hat{N}, \hat{\delta})$ is considered in the subsequent section.

2.1.4 The slope of the power function

Let (N, δ) be a test for hypothesis

$$H_o: \theta \leq \theta^* \quad \text{against} \quad H_1: \theta > \theta^*, \quad \underline{\theta} < \theta^* < \overline{\theta}, \tag{2.54}$$

with the power function $M(\theta)$, $\theta \in (\underline{\theta}, \overline{\theta})$. Then the slope of $M(\theta)$ at $\theta = \theta^*$ is a measure for the discriminatory power in the neighbourhood of θ^*. The subsequent lemma presents an expression for this slope if (N, δ) is a WLRT. Moreover, we shall consider a further method of the determination of the stopping bounds of a WLRT for hypotheses (2.54).

L e m m a 2.1.4. Let $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n \in \Gamma^+}$ be a closed WLRT, $\underline{\theta} < \theta_o < \theta_1 < \overline{\theta}$, based on a sequence of i.i.d. random variables having density

$$f_\theta(x) = h(x) \exp(d(\theta)t(x) - c(\theta)), \; x \in \mathfrak{X}, \; \theta \in (\underline{\theta}, \overline{\theta}). \tag{2.55}$$

Suppose that Lemma 1.6.4 holds. Further, we suppose that

$$P_\theta(L_{N,\theta_o,\theta_1} = B \mid H_o \text{ is accepted}) = 1 \tag{2.56}$$

and

$$P_\theta(L_{N,\theta_o,\theta_1} = A \mid H_1 \text{ is accepted}) = 1, \tag{2.57}$$

for $\theta \in (\underline{\theta}, \overline{\theta})$ and that the first derivative of $M(\theta)$ at $\theta = \theta^*$ exists, where θ^* denotes the separating-parameter given by (1.58).

Then

$$M'(\theta^*) = \left.\frac{dM(\theta)}{d\theta}\right|_{\theta=\theta^*} = - \frac{\ln A \ln B}{\ln A - \ln B} \frac{d'(\theta^*)}{d(\theta_1) - d(\theta_o)}. \tag{2.58}$$

P r o o f. Under the conditions of this lemma the power function is given by (2.30). This parametric form implies

$$\left.\frac{dM(\theta')}{d\theta'}\right|_{\theta'=\theta^*} = \left.\frac{dM^*(h)}{dh}\right|_{h=0} \cdot \left.\frac{dh}{d\theta'}\right|_{\theta'=\theta^*} \tag{2.59}$$

with $h = (d(\theta'') - d(\theta'))/(d(\theta_1) - d(\theta_o))$, where the corresponding θ'' is determined by $\zeta(\theta',\theta^*) = \zeta(\theta'',\theta^*)$. Since

$$M^*(h) = \frac{-\ln B}{\ln A - \ln B} + \frac{1}{2} \frac{\ln A \ln B}{\ln A - \ln B} h + o(h) \quad \text{for} \quad h \to 0$$

we obtain

$$\left.\frac{dM^*(h)}{dh}\right|_{h=0} = \frac{1}{2} \frac{\ln A \ln B}{\ln A - \ln B} \tag{2.60}$$

Now, if Lemma 1.6.4 holds, for every sequence $\{\theta'_n\}_{n\in\Gamma^+} \in (\underline{\theta},\theta^*)$ with $\lim_{n\to\infty} \theta'_n = \theta^*$ we obtain corresponding sequences $\{\theta''_n\}_{n\in\Gamma^+} \in (\theta^*,\bar{\theta})$ and $\{h_{\theta'_n}\}_{n\in\Gamma^+} \in (0, \infty)$, where θ''_n and $h_{\theta'_n}$ are determined by

$$\zeta(\theta'_n,\theta^*) = \zeta(\theta''_n,\theta^*)$$

and

$$h_{\theta'_n} = (d(\theta''_n) - d(\theta'_n))/(d(\theta_1) - d(\theta_o)),$$

so that

$$\lim_{n\to\infty} \theta''_n = \theta^* \quad \text{and} \quad \lim_{n\to\infty} h_{\theta'_n} = 0.$$

If instead of $\{\theta'_n\}_{n\in\Gamma^+}$ sequence $\{\theta''_n\}_{n\in\Gamma^+} \in (\theta^*,\bar{\theta})$ is given, a corresponding sequence $\{h_{\theta''_n}\}_{n\in\Gamma^+}$ exists so that

$$h_{\theta''_n} = (d(\theta'_n) - d(\theta''_n))/(d(\theta_1) - d(\theta_o)) = - h_{\theta'_n}, \quad n \in \Gamma^+.$$

Hence, we obtain

$$\left.\frac{dh}{d\theta'}\right|_{\theta'=\theta^*} = \lim_{\theta'_n \to \theta^*} \frac{h_{\theta''_n} - h_{\theta'_n}}{\theta''_n - \theta'_n}$$

$$= - \frac{2}{d(\theta_1) - d(\theta_o)} \lim_{\theta'_n \to \theta^*} \frac{d(\theta''_n) - d(\theta'_n)}{\theta''_n - \theta'_n}$$

$$= - \frac{2}{d(\theta_1) - d(\theta_o)} d'(\theta^*). \tag{2.61}$$

Collecting together (2.59), (2.60) and (2.61) we obtain (2.58).■

If (2.56) and (2.57) holds only approximately, the right-hand side of (2.58) provides an approximation for the slope of the power function at $\theta = \theta^*$.

E x a m p l e 2.1.3. Continuation of Example 2.1.0. We consider the slope of the power function at $\theta = \theta^*$ and assume that the excess of L_{N,θ_0,θ_1} at termination over B and A is small.

(i) The binomial proportion. We have $d(\theta) = \ln(\theta/(1-\theta))$ and the separating-parameter θ^* is determined by (1.79). Then, by Lemma 2.1.4 we obtain

$$M'(\theta^*) \approx -\frac{\ln A \ln B}{\ln A - \ln B}\left(\ln\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)^{-1}\frac{1}{\theta^*(1-\theta^*)}. \tag{2.62}$$

If we consider hypotheses

$$H_0:\ \theta = \theta_0 = \frac{1}{2} - \varepsilon \quad\text{and}\quad H_1:\ \theta = \theta_1 = \frac{1}{2} + \varepsilon\ ,\ 0 < \varepsilon < \frac{1}{2},$$

then we have $\theta^* = 1/2$, and (2.56) and (2.57) at least hold for some values of B and A (see Example 2.1.2). Under this additional assumption by (2.58) we obtain the exact value for the slope of $M(\theta)$ at $\theta^* = 1/2$:

$$M'(\tfrac{1}{2}) = -\frac{\ln A \ln B}{\ln A - \ln B}\,2\left(\ln\frac{1+2\varepsilon}{1-2\varepsilon}\right)^{-1} \tag{2.63}$$

For small values of ε we have $\ln((1+2\varepsilon)/(1-2\varepsilon)) \approx 4\varepsilon$ and (2.63) implies

$$M'(\tfrac{1}{2}) \approx -\frac{\ln A \ln B}{\ln A - \ln B}\,\frac{1}{2\varepsilon} \quad\text{for}\quad \varepsilon \to 0. \tag{2.64}$$

(ii) The Poisson mean. We have $d(\theta) = \ln\theta$ and θ^* is determined by (1.80). Then we obtain

$$M'(\theta^*) \approx -\frac{\ln A \ln B}{\ln A - \ln B}\,\frac{1}{\theta_1 - \theta_0}. \tag{2.65}$$

(iii) The normal mean. We have $d(\theta) = \theta/\sigma^2$ and θ^* is determined by (1.81). Then we obtain (2.65) again. We note that this approximation does not depend on variance σ^2.

(iv) The exponential mean. We suppose $f_\theta(x) = \theta\exp(-\theta x)$, $x \in (0,\infty)$, $\theta \in (0,\infty)$. Then we have $d(\theta) = -\theta$, θ^* is determined by (1.82), and again we obtain (2.65). ■

Now we shall return to the test $(\hat{N},\hat{\delta})$ of Section 2.1.3 for hypothesis $H_0:\ \theta \leq \theta^*$ against $H_1:\ \theta > \theta^*$ defined by (2.36), (2.37) and (2.38). Then the formula (2.58) can be modified as follows.

Corollary 2.1.3. Let $(\hat{N}, \hat{\delta})$ be a closed test for $H_o: \theta \leq \theta^*$ against $H_1: \theta > \theta^*$ defined by (2.36), (2.37) and (2.38). We suppose that

$$P_\theta(L_N^* = \hat{B} \mid H_o \text{ is accepted}) = 1 \tag{2.66}$$

and

$$P_\theta(L_N^* = \hat{A} \mid H_1 \text{ is accepted}) = 1 \tag{2.67}$$

for $\theta \in (\underline{\theta}, \bar{\theta})$ and that the first derivative of its power function $\hat{M}(\theta)$ exists for $\theta = \theta^*$. Then

$$\hat{M}'(\theta^*) = \left.\frac{d\hat{M}(\theta)}{d\theta}\right|_{\theta=\theta^*} = - \frac{\ln \hat{A} \ln \hat{B}}{\ln \hat{A} - \ln \hat{B}} d'(\theta^*). \tag{2.68}$$

Proof. It has been shown in the previous section that test $(\hat{N}, \hat{\delta})$ is identical with a WLRT $(N, \delta) = \{L_{N,\theta',\theta''}, \hat{B}^{d_{\theta',\theta''}}, \hat{A}^{d_{\theta',\theta''}}\}_{n \in \Gamma^+}$ for $H_o: \theta = \theta'$ against $H_1: \theta = \theta''$, where to any given θ', $\underline{\theta} < \theta' < \theta^*$, parameter θ'' is determined by (2.39), and $d_{\theta',\theta''}$ is given by (2.41). Denoting by $M(\theta)$ the power function of (N, δ) we obtain $M(\theta) = \hat{M}(\theta)$, $\theta \in (\underline{\theta}, \bar{\theta})$. Hence, applying Lemma 2.1.4 we obtain

$$\hat{M}'(\theta^*) = M'(\theta^*) = - \frac{\ln \hat{A}^{d_{\theta',\theta''}} \ln \hat{B}^{d_{\theta',\theta''}}}{\ln \hat{A}^{d_{\theta',\theta''}} - \ln \hat{B}^{d_{\theta',\theta''}}} \frac{d'(\theta^*)}{d_{\theta',\theta''}}$$

$$= - \frac{\ln \hat{A} \ln \hat{B}}{\ln \hat{A} - \ln \hat{B}} d'(\theta^*),$$

which completes the proof. ∎

Under the conditions of this corollary, the slope of the power function $\hat{M}(\theta)$ at $\theta = \theta^*$ of test $(\hat{N}, \hat{\delta})$ for $H_o: \theta \leq \theta^*$ against $H_1: \theta > \theta^*$ will depend beside $\hat{B}$ and $\hat{A}$ only on the given parameter θ^*. This property will provide a further possibility for the determination of the stopping bounds $\hat{B}$ and $\hat{A}$.

Corollary 2.1.4. We suppose that Corollary 2.1.3 holds. Let $d'(\theta^*) > 0$. If to given m^*, $0 < m^* < \infty$, the stopping bounds $\hat{B}$ and $\hat{A}$ are chosen by

$$\hat{B} = \exp(-2m^*/d'(\theta^*)) \tag{2.69}$$

and

$$\hat{A} = 1/\hat{B}, \tag{2.70}$$

then the power function $\hat{M}(\theta)$ of $(\hat{N}, \hat{\delta})$ satisfies

$$\hat{M}(\theta^*) = \frac{1}{2} \tag{2.71}$$

and

$$\hat{M}'(\theta^*) = m^*. \tag{2.72}$$

Proof. For $\theta' = \theta^*$, by (2.45) and (2.21), we obtain

$\hat{M}(\theta^*) = (-\ln \hat{B})/(\ln \hat{A} - \ln \hat{B})$ so that we have $\hat{M}(\theta^*) = 1/2$ iff $\hat{A} = 1/\hat{B}$. Hence, by (2.68) we obtain

$$\hat{M}'(\theta^*) = -\frac{\ln \hat{B}}{2} d'(\theta^*)$$

and requirement $\hat{M}'(\theta^*) = m^*$ is equivalent to (2.69).■

If instead of (2.66) and (2.67) we may only suppose that the excess of L_N^* at termination over $\hat{B}$ and $\hat{A}$ is small, then the right-hand sides of (2.69) and (2.70) provide approximations for $\hat{B}$ and $\hat{A}$ such that then $(\hat{N}, \hat{\delta})$ approximately satisfies requirements (2.71) and (2.72).

We remark, that sufficient conditions for the existence of the first derivative of the power function of a closed sequential test (N, δ) at $\theta = \theta^*$ have been considered by ABRAHAM [1] and BERK [13]. These conditions are fulfilled for the distributions considered here. Furthermore, ABRAHAM [1] has shown that

$$\left.\frac{dM(\theta)}{d\theta}\right|_{\theta=\theta^*} = E_{\theta^*} \delta S_N \quad \text{with} \quad S_n = \left.\frac{\partial \ln f_\theta(X_n)}{\partial \theta}\right|_{\theta=\theta^*}, \quad n \in \Gamma^+.$$

This equation is obtained formally from

$$\left.\frac{dM(\theta)}{d\theta}\right|_{\theta=\theta^*} = \frac{d}{d\theta} E_\theta \delta \chi_{\{N<\infty\}}\Big|_{\theta=\theta^*}$$

by differentiating across the expectation sign. We refer to [1].

2.1.5 Upper bounds for the true risks

We again consider WLRT $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$. For any given parameter pair $\theta', \theta'' \in \Theta$ with $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ and $h > 0$ let $\alpha(\theta')$ and $\beta(\theta'')$ be defined by

$$\alpha(\theta') = M(\theta') \tag{2.73}$$

and

$$\beta(\theta'') = Q(\theta''). \tag{2.74}$$

We shall denote $\alpha(\theta')$ and $\beta(\theta'')$ as the <u>true risks</u> of our test at $\theta = \theta'$ and $\theta = \theta''$, respectively. Especially, if $\theta' = \theta_0$ and $\theta'' = \theta_1$ holds we have $(\theta_0, \theta_1) \overset{h}{\sim} (\theta_0, \theta_1)$, and the true risks $\alpha(\theta_0)$ and $\beta(\theta_1)$ are the usual probabilities of an error of first and second kind.

<u>Lemma 2.1.5</u>. Let $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$ be a closed WLRT. If $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ with $h > 0$, then

$$\alpha(\theta') + B^{-h}\beta(\theta'') \leq 1 \tag{2.75}$$

and

$$A^h \alpha(\theta') + \beta(\theta'') \leq 1. \tag{2.76}$$

P r o o f. Since (N, δ) is closed, by Theorem 2.2.1, (2.73) and (2.74), we obtain

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1}\chi\{N<\infty\} \mid H_o \text{ is accepted}) = \beta(\theta'')/(1-\alpha(\theta')) \tag{2.77}$$

and

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1}\chi\{N<\infty\} \mid H_1 \text{ is accepted}) = (1-\beta(\theta''))/\alpha(\theta'). \tag{2.78}$$

Otherwise, since $0<B<1<A<\infty$ and $h>0$ we obtain

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1}\chi\{N<\infty\} \mid H_o \text{ is accepted}) \leqslant B^h \tag{2.79}$$

and

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1}\chi\{N<\infty\} \mid H_1 \text{ is accepted}) \geqslant A^h. \tag{2.80}$$

Collecting together (2.77) until (2.80) we obtain (2.75) and (2.76).■

The consequences of inequalities (2.75) and (2.76) are shown in Fig. 2.3. The true risks must belong to the shaded quatriliteral.

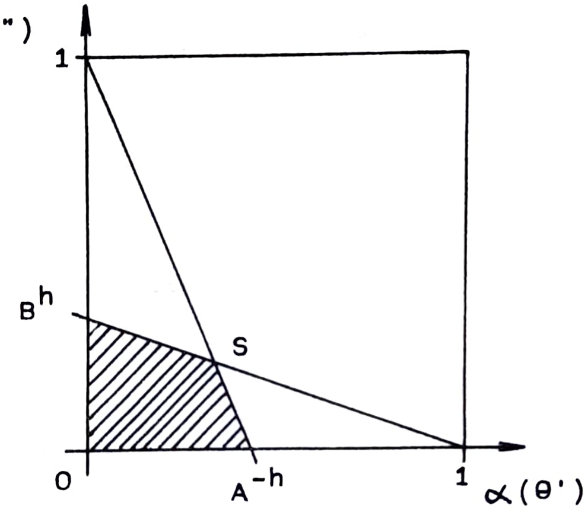

Fig. 2.3 Possible domain of the true risks $\alpha(\theta')$ and $\beta(\theta'')$

This implies the quantities A^{-h} and B^h are upper bounds for the true risks $\alpha(\theta')$ and $\beta(\theta'')$ and we have

$$\alpha(\theta') \leqslant A^{-h} \quad \text{and} \quad \beta(\theta'') \leqslant B^h \tag{2.81}$$

in each case. Especially, since $(\theta_o,\theta_1) \overset{1}{\sim} (\theta_o,\theta_1)$ we obtain

$$\alpha(\theta_o) \leqslant A^{-1} \quad \text{and} \quad \beta(\theta_1) \geqslant B.$$

That means, to given α and β, $0<\alpha,\beta<1$, the choice

$$A = \alpha^{-1} \quad \text{and} \quad B = \beta \tag{2.82}$$

is sufficient to obtain a test with $M(\theta_o) \leqslant \alpha$ and $M(\theta_1) \geqslant 1-\beta$.

We notice that for $0<B<1<A<\infty$ and $h>0$ by (2.77) until (2.80) we still obtain the inequality

$$\frac{\beta(\theta'')}{1-\alpha(\theta')} < \frac{1-\beta(\theta'')}{\alpha(\theta')}, \tag{2.83}$$

which is equivalent to

$$\alpha(\theta') + \beta(\theta'') < 1.$$

Therefore, line $\alpha(\theta') + \beta(\theta'') = 1$ is an absolute upper bound to the possible domain of the true risks of our test.

2.1.6 Bounds for the power function

The bounds for the true risks of a WLRT considered in the previous section can be used to obtain two-sided bounds for the power function.

Lemma 2.1.6. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ a closed WLRT. Then for every $\theta' \in \Theta$ with $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ we have

$$\max\{0, 1-B^{-h}\} \leq M(\theta') \leq \min\{1, A^{-1}\}. \tag{2.84}$$

Proof. By $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h>0$ and (2.81) we obtain

$$M(\theta') \leq A^{-h} \qquad \text{and} \qquad M(\theta'') \geq 1 - B^{h}.$$

If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h<0$, we obtain $(\theta'',\theta') \overset{-h}{\sim} (\theta_0,\theta_1)$ with $-h>0$. Applying (2.81) again, we obtain

$$M(\theta'') \leq A^{-h} \qquad \text{and} \qquad M(\theta') \geq 1 - B^{-h}.$$

Furthermore, we have $0 \leq M(\theta') \leq 1$ for $\theta' \in \Theta$. Collecting together these inequalities, we obtain (2.84).■

Inequality (2.84) can be improved if additional assumptions are fulfilled. Two cases may be of special interest.

(i) We suppose $P_{\theta'}(L_{N,\theta_0,\theta_1} = B \mid H_0 \text{ is accepted}) = 1$ for $\theta' \in \Theta$: That means we assume that only the excess of L_{N,θ_0,θ_1} at termination over B is zero. For instance, such an assumption can be fulfilled for certain WLRTs concerning the binomial proportion or the Poisson mean. We refer to Example 2.1.4. Then we obtain the following two-sided bounds for the power function.

Lemma 2.1.7. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT where

$$P_{\theta'}(L_{N,\theta_0,\theta_1} = B \mid H_0 \text{ is accepted}) = 1 \text{ for } \theta' \in \Theta. \tag{2.85}$$

Then for every $\theta' \in \Theta$ with $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ we have

$$\max\{0, 1-B^{-h}\} \leq M(\theta') \leq M^{*}(h) \tag{2.86}$$

where M^* is defined by (2.21).

P r o o f. By (2.12), (2.73), (2.74) and (2.85) we obtain

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1}\chi_{\{N<\infty\}}|H_0 \text{ is accepted}) = \beta(\theta'')/(1-\alpha(\theta'))$$
$$= B^h,$$

respectively

$$\alpha(\theta') + B^{-h}\beta(\theta'') = 1 \tag{2.87}$$

so that point $(\alpha(\theta'),\beta(\theta''))$ belongs to the finite line between the intersection point $(0,B^h)$ of the straight line, given by equation (2.87), with the $\beta(\theta'')$-axis and intersection point

$$S = \left(\frac{1-B^h}{A^h-B^h},\ B^h\frac{A^h-1}{A^h-B^h}\right) \tag{2.88}$$

of the straight lines given by equations (2.87) and (2.76). Compare Fig. 2.3. For $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h>0$ this implies

$$0 \leq \alpha(\theta') \leq (1-B^h)/(A^h-B^h) = M^*(h)$$

and

$$M^*(-h) = B^h(A^h-1)/(A^h-B^h) \leq \beta(\theta'') \leq B^h$$

or

$$0 \leq M(\theta') \leq M^*(h)$$

and

$$1 - B^h \leq M(\theta'') \leq 1 - B^h(A^h-1)/(A^h-B^h) = M^*(-h),$$

respectively. Analogously, for $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h<0$, we obtain

$$0 \leq M(\theta'') \leq M^*(-h)$$

and

$$1 - B^{-h} \leq M(\theta') \leq M^*(h).$$

Collecting together these inequalities, we obtain (2.86).■

Hence, if (2.85) is true the WALD approximation $M^*(h)$ for the power function $M(\theta)$ is an upper bound for $M(\theta)$ if $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$. An analogous result that contains a corresponding lower bound for $M(\theta)$ can be obtained if $P_{\theta'}(L_{N,\theta_0,\theta_1} = A\,|\,H_1 \text{ is accepted}) = 1$.

Now, inequality (2.86) can be used to obtain values for the stopping bounds B and A of $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ so that $M(\theta_0) \leq \alpha$ and $M(\theta_1) \geq 1-\beta$ holds to given α and β, $0<\alpha,\beta<1$.

C o r o l l a r y 2.1.5. We suppose that Lemma 2.1.7 holds. If to given α and β, $0<\alpha,\beta<1$, $\alpha+\beta<1$, the stopping bounds B and A are chosen by

$$B = \beta \quad\text{and}\quad A = \frac{1-\beta+\alpha\beta}{\alpha}, \tag{2.89}$$

then we have $\quad M(\theta_0) \leq \alpha \quad\text{and}\quad M(\theta_1) \geq 1-\beta.$ $\qquad$ (2.90)

P r o o f. Since $(\theta_o,\theta_1) \overset{1}{\sim} (\theta_o,\theta_1)$ by (2.86) we obtain

$$M(\theta_o) \leq M^*(1) = (1 - B)/(A - B)$$

and

$$M(\theta_1) \geq 1 - B.$$

If we put $(1 - B)/(A - B) = \alpha$ and $1 - B = 1 - \beta$ which is equivalent to (2.89), we obtain (2.90). ■

We consider an example, where assumption (2.85) holds at least for some values of B.

<u>E x a m p l e 2.1.4</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ be a closed WLRT based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density

$$f_\theta(x) = h(x)\exp(d(\theta)\cdot x - c(\theta)),\ x\in\Gamma_o^+,\ \theta\in(\underline{\theta},\overline{\theta}).$$

Then we have

$$L_{n,\theta_o,\theta_1} = \exp\ (d(\theta_1) - d(\theta_o))\sum_{i=1}^{n}X_i - n(c(\theta_1) - c(\theta_o))$$

and

$$Z_{n,\theta_o,\theta_1} = (d(\theta_1) - d(\theta_o))\sum_{i=1}^{n}X_i - n(c(\theta_1) - c(\theta_o)),$$

$n\in\Gamma^+$. We suppose that an integer $g\in\Gamma^+$ exists so that

$$d(\theta_1) - d(\theta_o) = g(c(\theta_1) - c(\theta_o)). \tag{2.91}$$

This condition may be fulfilled, for instance, for special WLRTs concerning the mean of a Bernoulli or Poisson distribution. Then

$$Z_{n,\theta_o,\theta_1} = (g\sum_{i=1}^{n}X_i - n)(c(\theta_1) - c(\theta_o)),\ n\in\Gamma^+.$$

and Z_{n,θ_o,θ_1} is an integer multiple of $c(\theta_1) - c(\theta_o)$ for every $n\in\Gamma^+$. If we further may suppose that

$$c(\theta_1) - c(\theta_o) > 0, \tag{2.92}$$

then for every B, $0 < B < 1$, an integer $k_B\in\Gamma^+$ exists with

$$k_B = \max\{k\in\Gamma^+ :\ -k(c(\theta_1) - c(\theta_o)) \leq \ln B\}.$$

and we have

$$Z_{n,\theta_o,\theta_1} = -k_B(c(\theta_1) - c(\theta_o)) \text{ on } \{N = n\}\cap\{H_o \text{ is accepted}\}$$

$n\in\Gamma^+$. Hence, we obtain

$$P_{\theta'}(L_{N,\theta_o,\theta_1} = \exp(-k_B(c(\theta_1) - c(\theta_o))|\,H_o \text{ is accepted}) = 1$$

so that Lemma 2.1.7 and Corollary 2.1.5 is applicable if

$$B = \exp(-k_B(c(\theta_1) - c(\theta_0))).$$

We remark that under conditions (2.91) and (2.92) test $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ is identical with test

$$(N,\delta) = \{L_{n,\theta_0,\theta_1}, \exp(-k_B(c(\theta_1) - c(\theta_0))), A\}_{n\in\Gamma^+}. \blacksquare$$

(ii) The symmetrical case: We consider test $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ and suppose that

$$B = 1/A \tag{2.93}$$

and

$$\alpha(\theta') = \beta(\theta'') \quad \text{for} \quad (\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1). \tag{2.94}$$

Then we obtain the following bounds for the power function.

<u>L e m m a 2.1.8</u>. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT, where (2.93) and (2.94) hold. If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$, then

$$M(\theta') \leqslant M^*(h) \quad \text{for} \quad h > 0 \tag{2.95}$$

and

$$M(\theta') \leqslant M^*(h) \quad \text{for} \quad h < 0 \tag{2.96}$$

where M^* is defined by (2.21).

P r o o f. If (2.94) holds, then point $(\alpha(\theta'), \beta(\theta''))$ is a point on the finite line given by points (0,0) and S, where S is defined by (2.88). This, $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$ and (2.93) imply

$$0 \leqslant \alpha(\theta') = M(\theta') \leqslant B^h/(1 + B^h) = M^*(h) \tag{2.97}$$

and

$$0 \leqslant \beta(\theta'') = 1 - M(\theta'') \leqslant B^h/(1 + B^h) = M^*(h), \tag{2.98}$$

where (2.97) establishes (2.95). If $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h < 0$, then we obtain $(\theta'',\theta') \overset{-h}{\sim} (\theta_0,\theta_1)$ with $-h > 0$. Applying (2.97) and (2.98), we obtain now

$$0 \leqslant \alpha(\theta'') = M(\theta'') \leqslant B^{-h}/(1 + B^{-h}) = M^*(-h)$$

and

$$0 \leqslant \beta(\theta') = 1 - M(\theta') \leqslant B^{-h}/(1 + B^{-h}) = M^*(-h).$$

This and $M^*(h) + M^*(-h) = 1$ provide (2.96). $\blacksquare$

Hence, under the conditions of this lemma the WALD approximation $M^*(h)$ for $M(\theta')$ is an upper bound for $M(\theta')$ if $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$ and a lower bound for $M(\theta')$ if $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h < 0$, respectively. This property can be used to obtain a WLRT that is admissible in the following sense.

<u>C o r o l l a r y 2.1.6</u>. Suppose that Lemma 2.1.8 holds. If to given $\alpha < 1/2$ the stopping bounds B and A of (N,δ) are chosen by

given $\alpha < 1/2$ the stopping bounds B and A of (N, δ) are chosen by

$$B = \alpha/(1-\alpha) \quad \text{and} \quad A = 1/B, \tag{2.99}$$

then we have

$$M(\theta_0) \leq \alpha \quad \text{and} \quad M(\theta_1) \geq 1 - \alpha . \tag{2.100}$$

P r o o f. Applying Lemma 2.1.8 it follows immediately from

$$M(\theta_0) \leq M^*(1) = B/(1-B) \leq \alpha . \blacksquare$$

We note that for the symmetrical case considered here the stopping bounds (2.99) proposed by this corollary coincide with the corresponding WALD approximations. An example for this symmetrical case is a WLRT for the normal mean with known variance and equal probabilities of an error of first and second kind, respectively.

2.2 Most powerful tests

Let (N, δ) be a test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. Then the properties if its power function $M(\theta)$ depend on sample size N as well as terminal decision rule δ. Here we investigate a possibility to exert an influence on the power function by a suitable choice of the terminal decision rule if sample size N is given. We shall see that special LRTs with quite a simple structure of the terminal decision rule are best tests in the sense of the following definition.

D e f i n i t i o n 2.2.1. (i) A test (N, δ) for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ is said to be a test at (significance) level α, $0 < \alpha < 1$, iff

$$M(\theta_0) = E_{\theta_0} \delta \chi_{\{N < \infty\}} . \tag{2.101}$$

(ii) Let $\mathcal{T}_\alpha(N)$ be the set of all tests (N, δ) for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ at level α at sample size N. A test $(N, \hat{\delta}) \in \mathcal{T}_\alpha(N)$ is said to be a most powerful test (MP-test) at level α iff

$$\hat{M}(\theta_1) = E_{\theta_1} \hat{\delta} \chi_{\{N < \infty\}} = \sup_{(N, \delta) \in \mathcal{T}_\alpha(N)} E_{\theta_1} \delta \chi_{\{N < \infty\}} . \tag{2.102}$$

According to this definition a most powerful test $(N, \hat{\delta})$ at sample size N maximizes the probability of acceptance of the hypothesis $H_1: \theta = \theta_1$ for $\theta = \theta_1$ within the class of all test (N, δ) at level α. The subsequent theorem shows that the well-known Lemma of NEYMAN and PEARSON (see e.g. [53]) which characterizes the structure of the decision rule of a most powerful fixed-sample test can be extended to sequential tests.

Theorem 2.2.1. Let $(\Omega, \mathcal{F}, \mathcal{P} = \{P_\theta, \theta \in \Theta\})$ be a statistical structure such that a non-decreasing sequence $\{\mathcal{F}_n\}_{n\in\Gamma^+}$ of sub-σ-algebras of $\mathcal{F}$ and to given $\theta_o, \theta_1 \in \Theta$, $\theta_o \neq \theta_1$, the corresponding sequence $\{L_{n,\theta_o,\theta_1}\}_{n\in\Gamma^+}$ of likelihood ratios exist. Let N be a stopping time w.r.t. $\{\mathcal{F}_n\}_{n\in\Gamma^+}$ with

$$P_{\theta_o}(N < \infty) = \alpha^* > 0 \tag{2.103}$$

Then for every α, $0 < \alpha < \alpha^*$, an MP-test $(N, \hat{\delta})$ at level α for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ exists, whose terminal decision rule $\hat{\delta}$ is given by

$$\hat{\delta} = \chi_{\{L_{N,\theta_o,\theta_1} > c_\alpha\}} + \gamma_\alpha \chi_{\{L_{N,\theta_o,\theta_1} = c_\alpha\}}, \tag{2.104}$$

where c_α, $0 \leq c_\alpha < \infty$, and γ_α, $0 \leq \gamma_\alpha \leq 1$, are constants determined by

$$P_{\theta_o}(L_{N,\theta_o,\theta_1} > c_\alpha, N < \infty) + \gamma_\alpha P_{\theta_o}(L_{N,\theta_o,\theta_1} = c_\alpha, N < \infty) = \alpha. \tag{2.105}$$

Proof. We consider the probability $P_{\theta_o}(L_{N,\theta_o,\theta_1} \geq z, N < \infty)$ as a function of z, $0 \leq z < \infty$. This probability is a non-increasing function of z with $P_{\theta_o}(L_{N,\theta_o,\theta_1} \geq 0, N < \infty) = P_{\theta_o}(N < \infty) = \alpha^*$, so that for any given α, $0 < \alpha < \alpha^*$, real numbers c_α, $0 \leq c_\alpha < \infty$, and γ_α, $0 \leq \gamma_\alpha \leq 1$, exist with

$$P_{\theta_o}(L_{N,\theta_o,\theta_1} > c_\alpha, N < \infty) + \gamma_\alpha P_{\theta_o}(L_{N,\theta_o,\theta_1} = c_\alpha, N < \infty) = \alpha. \tag{2.106}$$

Then, for the power function $\hat{M}(\theta)$ of $(N, \hat{\delta})$ we have

$$\begin{aligned}
\hat{M}(\theta_o) &= E_{\theta_o} \hat{\delta} \chi_{\{N<\infty\}} \\
&= E_{\theta_o} \chi_{\{L_{N,\theta_o,\theta_1} > c_\alpha, N<\infty\}} + \gamma_\alpha E_{\theta_o} \chi_{\{L_{N,\theta_o,\theta_1} = c_\alpha, N<\infty\}} \\
&= P_{\theta_o}(L_{N,\theta_o,\theta_1} > c_\alpha, N < \infty) + \gamma_\alpha P_{\theta_o}(L_{N,\theta_o,\theta_1} = c_\alpha, N < \infty) \\
&= \alpha
\end{aligned}$$

and $(N, \hat{\delta})$ is a test at level α. Let now (N, δ) be any arbitrary test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at level α with the power function $M(\theta)$. Let Ω^+ be the set defined by

$$\Omega^+ = \{\omega : \hat{\delta} > \delta, N < \infty\}. \tag{2.107}$$

Since $0 \leq \delta \leq 1$ we have $\hat{\delta} > 0$ on Ω^+ and therefore $L_{N,\theta_o,\theta_1} \geq c_\alpha$ on Ω^+. This implies

$$(\hat{\delta} - \delta)(L_{N,\theta_o,\theta_1} - c_\alpha) \geq 0 \quad \text{on} \quad \Omega^+. \tag{2.108}$$

Let Ω^- be defined by

$$\Omega^- = \{\hat{\delta} < \delta, N < \infty\}. \tag{2.109}$$

Since $0 \le \delta \le 1$ we have $\hat{\delta} < 1$ and therefore $L_{N,\theta_0,\theta_1} \le c_\alpha$ on Ω^-. This also implies

$$(\hat{\delta} - \delta)(L_{N,\theta_0,\theta_1} - c_\alpha) \ge 0 \quad \text{on} \quad \Omega^-. \tag{2.110}$$

Putting together (2.108) and (2.110) we obtain

$$(\hat{\delta} - \delta)(L_{N,\theta_0,\theta_1} - c_\alpha) \ge 0 \quad \text{on} \quad \Omega^+ \cup \Omega^-. \tag{2.111}$$

This implies

$$E_{\theta_0}((\hat{\delta} - \delta)(L_{N,\theta_0,\theta_1} - c_\alpha)\chi_{\{N<\infty\}}) \ge 0. \tag{2.112}$$

Otherwise, we have

$$E_{\theta_0}((\hat{\delta} - \delta)(L_{N,\theta_0,\theta_1} - c_\alpha)\chi_{\{N<\infty\}})$$

$$= E_{\theta_0}(\hat{\delta} - \delta)L_{N,\theta_0,\theta_1}\chi_{\{N<\infty\}} - c_\alpha E_{\theta_0}(\hat{\delta} - \delta)\chi_{\{N<\infty\}}. \tag{2.113}$$

Since $E_{\theta_0}\hat{\delta}\chi_{\{N<\infty\}} = \hat{M}(\theta_0) = \alpha$ and (N, δ) is a test at level α, we obtain

$$E_{\theta_0}(\hat{\delta} - \delta)\chi_{\{N<\infty\}} = \hat{M}(\theta_0) - M(\theta_0) \ge 0. \tag{2.114}$$

Further, applying Lemma 1.6.1 we obtain

$$E_{\theta_0}(\hat{\delta} - \delta)L_{N,\theta_0,\theta_1}\chi_{\{N<\infty\}} = E_{\theta_1}\hat{\delta}\chi_{\{N<\infty\}} - E_{\theta_1}\delta\chi_{\{N<\infty\}}$$

$$= \hat{M}(\theta_1) - M(\theta_1). \tag{2.115}$$

Collecting together (2.112) to (2.115) we obtain $\hat{M}(\theta_1) \ge M(\theta_1)$ and $(N, \hat{\delta})$ is an MP-test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ at level α. ■

We remark that the terminal decision rule $\hat{\delta}$ of an MP-test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ proposed by this theorem can also be written in the form

$$\hat{\delta} = \sum_{n \in \bar{\Gamma}^+} \hat{\delta}_n \chi_{\{N=n\}}, \tag{2.116}$$

where $\hat{\delta}_n$ is defined by

$$\hat{\delta}_n = \chi_{\{L_{n,\theta_0,\theta_1} > c_\alpha\}} + \gamma_\alpha \chi_{\{L_{n,\theta_0,\theta_1} = c_\alpha\}}, \quad n \in \bar{\Gamma}^+, \tag{2.117}$$

Particularly, like in the non-sequential case if $P_{\theta_0}(L_{N,\theta_0,\theta_1} = c) = 0$ for every c, $0 \le c < \infty$, the MP-test $(N, \hat{\delta})$ for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ has a non-randomized terminal decision rule $\hat{\delta}$. If, on the other hand, L_{N,θ_0,θ_1} is a discrete random variable, the terminal decision rule of the MP-test is strictly randomized at least for some

values of α. Of course, direct computation of the constants c_α and γ_α is difficult and will depend on the possibilities of obtaining corresponding assertions on the distribution of L_{N,θ_0,θ_1}. We consider the following example.

<u>E x a m p l e 2.2.1</u>. The uniform distribution. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having density

$$f_\theta(x) = \left(\frac{1}{\theta}\right)^{\chi\{x\in[0,\theta]\}} 0^{\chi\{x\notin[0,\theta]\}}, \quad x\in R^1,\ \theta\in(0,\infty).$$

Consider a test for hypothesis

$$H_0: \theta = \theta_0 \quad \text{against} \quad H_1: \theta = \theta_1, \quad 0<\theta_1<\theta_0<\infty.$$

Then we obtain

$$L_{1,\theta_0,\theta_1} = (\theta_0/\theta_1)^{\chi\{X_1\in[0,\theta_1]\}}\, 0^{\chi\{X_1\notin[0,\theta_1]\}}$$

and because of the $\{X_n\}_{n\in\Gamma^+}$ are i.i.d.

$$L_{n,\theta_0,\theta_1} = \prod_{i=1}^{n} (\theta_0/\theta_1)^{\chi\{X_i\in[0,\theta_1]\}}\, 0^{\chi\{X_i\notin[0,\theta_1]\}}$$

$$= (\theta_0/\theta_1)^{n\chi\{\max\{X_1,\ldots,X_n\}\le\theta_1\}}\, 0^{(1-\chi\{\max\{X_1,\ldots,X_n\}\le\theta_1\})}$$

$n\in\Gamma^+$. We remark that this example is one of the few examples where the likelihood ratios L_{n,θ_0,θ_1} turn out to be discrete, even though the $\{X_n\}_{n\in\Gamma^+}$ are continuous.

Now consider a test $(N, \hat{\delta})$ where N and $\hat{\delta}$ to given $a>0$ are defined by

$$N = \begin{cases} \inf\{n\geq 1: L_{N,\theta_0,\theta_1}\notin(0,a)\}, & \text{if such an } n \text{ exists,} \\ \infty, & \text{otherwise,} \end{cases}$$

and

$$\hat{\delta} = \chi_{\{L_{N,\theta_0,\theta_1}>c_\alpha, N<\infty\}} + \gamma_\alpha \chi_{\{L_{N,\theta_0,\theta_1} = c_\alpha, N<\infty\}},$$

where c_α and γ_α are constants still to be determined. According to Theorem 2.2.1 this test possesses the structure of an MP-test. We stop sampling and decide for H_1 if L_{n,θ_0,θ_1} is sufficiently large, which is an indication that probably hypothesis H_1 is true. Otherwise, we stop sampling and decide for H_0 if $L_{n,\theta_0,\theta_1} = 0$ because then it will be evident that H_0 is true.

We remark that

$$\lim_{n\to\infty} L_{n,\theta_0,\theta_1} = \infty$$

if H_1 is true so that bound a is always overcrossed if H_1 is true.

Moreover, it is evident that $P_{\theta_o}(N < \infty) = 1$ holds.

In order to determine the quantities c_α and γ_α to given α, $0 < \alpha < 1$, we consider the distribution of L_{N,θ_o,θ_1}. Without any loss of generality we may suppose that an integer $n^* \geq 1$ exists, where

$$(\theta_o/\theta_1)^{n^*} = a.$$

Then L_{N,θ_o,θ_1} takes on only the two values 0 and a, and we have

$$P_{\theta_o}(L_{N,\theta_o,\theta_1} = a) = P_{\theta_o}\left(\bigcap_{i=1}^{n^*}\{X_i \leq \theta_1\}\right) = (\theta_1/\theta_o)^{n^*}$$

and

$$P_{\theta_o}(L_{N,\theta_o,\theta_1} = 0) = 1 - (\theta_1/\theta_o)^{n^*}.$$

We distinguish between the following three cases:

(i) $P_{\theta_o}(L_{N,\theta_o,\theta_1} = a) = \alpha$: Then we accept H_1 with probability α if H_o is true iff

$$0 < c_\alpha < a \quad \text{and} \quad \gamma_\alpha = 0.$$

Moreover, we have $\hat{M}(\theta_1) = E_{\theta_1}\hat{\delta}\chi_{\{N<\infty\}} = 1$. That means we accept H_1 with probability one if H_1 is true. Hence, test $(N, \hat{\delta})$ is a so-called <u>power one test</u> at level α, which is an MP-test obviously.

(ii) $P_{\theta_o}(L_{N,\theta_o,\theta_1} = a) < \alpha$: Then, the acceptance of H_1 is not sufficient only in case of L_{N,θ_o,θ_1} to reach significance level α. Therefore, we have also to accept H_1 with a certain probability if $L_{N,\theta_o,\theta_1} = 0$. Hence, we choose $c_\alpha = 0$ and obtain

$$\alpha = \hat{M}(\theta_o) = P_{\theta_o}(L_{N,\theta_o,\theta_1} > 0) + \gamma_\alpha P_{\theta_o}(L_{N,\theta_o,\theta_1} = 0)$$

$$= (\theta_1/\theta_o)^{n^*} + \gamma_\alpha(1 - (\theta_1/\theta_o)^{n^*})$$

iff

$$\gamma_\alpha = (\alpha - (\theta_1/\theta_o)^{n^*})/(1 - (\theta_1/\theta_o)^{n^*}).$$

Again $(N, \hat{\delta})$ is a power one test.

(iii) $P_{\theta_o}(L_{N,\theta_o,\theta_1} = a) > \alpha$: Here we have to reject H_1 with a certain probability, even in case of $L_{N,\theta_o,\theta_1} = a$ to reach significance level α. Thus, we choose $c_\alpha = a$ and obtain

$$\alpha = \hat{M}(\theta_o) = P_{\theta_o}(L_{N,\theta_o,\theta_1} > a) + \gamma_\alpha P_{\theta_o}(L_{N,\theta_o,\theta_1} = a)$$

$$= 0 + \gamma_\alpha(\theta_1/\theta_o)^{n^*}$$

iff

$$\gamma_\alpha = \alpha/(\theta_1/\theta_o)^{n^*}.$$

Then, we only have $\hat{M}(\theta_1) = \gamma_\alpha$. This case shows that to a given sample size the requirement for the implementation of a given significance level may be invalid because a formal implementation of the given significance level may lead to obviously false decisions. ■

If the parameter space Θ consists of more than two parameters, the assertion of Theorem 2.2.1 can be extended as follows.

Corollary 2.2.1. Let $(N, \hat{\delta})$ be an MP-test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ at level α according to Theorem 2.2.1 with the power function $\hat{M}(\theta)$. Let (N, δ) be any other test for these hypotheses with power function $M(\theta)$.

(i) Let Θ' and Θ'' be non-empty subsets of Θ defined by

$$\Theta' = \{\theta' \in \Theta : \text{There exists a } \theta'' \in \Theta \text{ with } (\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1) \text{ and } h > 0\}$$

and

(2.118)

$$\Theta'' = \{\theta'' \in \Theta : (\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1),\ \theta' \in \Theta'\}. \tag{2.119}$$

Then

$$\hat{M}(\theta) \geq M(\theta) \text{ for } \theta \in \Theta' \text{ implies } \hat{M}(\theta) \geq M(\theta) \text{ for } \theta \in \Theta''. \tag{2.120}$$

(ii) Let $\hat{\Theta}$ and $\hat{\hat{\Theta}}$ be subsets of Θ defined by

$$\hat{\hat{\Theta}} = \{\hat{\hat{\theta}} \in \Theta : \text{There exists a } \hat{\theta} \in \Theta \text{ with } (\hat{\theta}, \hat{\hat{\theta}}) \overset{h}{\sim} (\theta_0, \theta_1) \text{ and } h \geq 0\}$$

and

(2.122)

$$\hat{\Theta} = \{\hat{\theta} \in \Theta : (\hat{\theta}, \hat{\hat{\theta}}) \overset{h}{\sim} (\theta_0, \theta_1),\ \hat{\hat{\theta}} \in \hat{\hat{\Theta}}\}. \tag{2.123}$$

Then

$$\hat{M}(\theta) \leq M(\theta) \text{ for } \theta \in \hat{\hat{\Theta}} \text{ implies } \hat{M}(\theta) \leq M(\theta) \text{ for } \theta \in \hat{\Theta}. \tag{2.124}$$

Proof. First, we remark that because of $(\theta_0, \theta_1) \overset{1}{\sim} (\theta_0, \theta_1)$ the introduced sets Θ', Θ'', $\hat{\Theta}$ and $\hat{\hat{\Theta}}$ are non-empty sets. Further, it has been shown in the proof of Theorem 2.2.1, cf. (2.111), that

$$(\hat{\delta} - \delta)(L_{N,\theta_0,\theta_1} - c_\alpha) \geq 0 \text{ on } \Omega^+ \cup \Omega^-, \tag{2.125}$$

where Ω^+ and Ω^- are defined by (2.107) and (2.108), respectively. Then, (2.125) implies

$$E_\theta(\hat{\delta} - \delta)(L^h_{N,\theta_0,\theta_1} - c^h_\alpha)\chi_{\{N<\infty\}} \geq 0 \tag{2.126}$$

and

$$E_\theta(\hat{\delta} - \delta)(L^{-h}_{N,\theta_0,\theta_1} - c^{-h}_\alpha)\chi_{\{N<\infty\}} \leq 0 \tag{2.127}$$

for $h > 0$ and $\theta \in \Theta$.

(i) For any given $\theta'' \in \Theta''$ let $\theta' \in \Theta$ be the corresponding $\theta' \in \Theta'$ with $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ and $h > 0$. Then, by (2.126) and Lemma 1.6.1 we obtain

$$E_{\theta'}(\hat{\delta} - \delta)(L^h_{N,\theta_0,\theta_1} - c^h_\alpha)\chi_{\{N<\infty\}}$$

$$= E_{\theta'}(\hat{\delta} - \delta) L^h_{N,\theta_0,\theta_1} \chi\{N<\infty\} - c^h_\alpha E_{\theta'}(\hat{\delta} - \delta)\chi\{N<\infty\}$$

$$= E_{\theta''}(\hat{\delta} - \delta)\chi\{N<\infty\} - c^h_\alpha E_{\theta'}(\hat{\delta} - \delta)\chi\{N<\infty\}$$

$$\geqslant 0. \tag{2.128}$$

Since $\hat{M}(\theta') \geqslant M(\theta')$ for $\theta' \in \Theta'$, this implies $\hat{M}(\theta'') \geqslant M(\theta'')$ for $\theta'' \in \Theta''$.
(ii) Assertion (2.124) is established analogously by means of (2.127). ■

The property of an MP-test characterized by this corollary can be interpreted as a restricted uniformly most powerful property. On condition that this corollary holds every uniform reduction of the probability of acceptance of H_1 for $\theta \in \Theta'$ by applying a terminal decision rule δ which differs from $\hat{\delta}$ effects a uniform reduction of this probability for $\theta \in \Theta''$. Conversely, a corresponding uniform improvement of the power function for $\theta \in \hat{\hat{\Theta}}$ effects a uniform increase of the power function for $\theta \in \hat{\Theta}$.
We consider the following example.

E x a m p l e 2.2.2. Let $(N, \hat{\delta})$ be an MP-test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, $\underline{\theta} < \theta_0 < \theta_1 < \bar{\theta}$, according to Theorem 2.2.1 based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density

$$f_\theta(x) = h(x) \exp(d(\theta)t(x) - c(\theta)), \quad x \in \mathfrak{X}, \quad \theta \in (\underline{\theta}, \bar{\theta}).$$

Suppose that Lemma 1.6.4 holds. Let θ^* be the separating-parameter given by (1.58). Then, by Lemma 1.6.4 to each $\theta' < \theta^*$ a $\theta'' > \theta^*$ corresponds so that

$$\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*). \tag{2.129}$$

This correspondence is a one-to-one correspondence between the elements of $(\underline{\theta}, \theta^*)$ and the elements of $(\theta^*, \bar{\theta})$. Applying Lemma 1.6.5, we obtain

$$(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1) \quad \text{with} \quad h = \frac{d(\theta'') - d(\theta')}{d(\theta_1) - d(\theta_0)} > 0$$

for every pair $\theta', \theta'' \in (\underline{\theta}, \bar{\theta})$ which satisfies (2.129). Then a possible choice for Θ' is $\Theta' = (\underline{\theta}, \theta^*)$, and we obtain

$$\Theta'' = \{\theta'' \in (\underline{\theta}, \bar{\theta}) : (\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1),\ \theta' \in (\underline{\theta}, \theta^*)\} = (\theta^*, \bar{\theta}).$$

We remark that a possible choice of sets $\hat{\Theta}$ and $\hat{\hat{\Theta}}$ considered in the second part of Corollary 2.2.1 is again

$$\hat{\Theta} = (\underline{\theta}, \theta^*) \quad \text{and} \quad \hat{\hat{\Theta}} = (\theta^*, \bar{\theta}).$$

This can be shown analogously.
If now (N, δ) is a further test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ with

$$M(\theta) \leq \hat{M}(\theta) \quad \text{for} \quad \theta \in (\underline{\theta}, \theta'') \cup \{\theta''\}, \tag{2.130}$$

then Corollary 2.2.1 also implies

$$M(\theta) \leq \hat{M}(\theta) \quad \text{for} \quad \theta \in (\theta^*, \bar{\theta}).$$

Hence, already (2.130) ensures here that $M(\theta)$ is not greater than $\hat{M}(\theta)$ for every $\theta \in (\underline{\theta}, \bar{\theta})$. For special examples satisfying the above assumptions we refer to the distributions considered in Example 1.6.1. ■

We present a further consequence of Theorem 2.2.1 and Corollary 2.2.1.

<u>Corollary 2.2.2</u>. Let $(N, \hat{\delta})$ be an MP-test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ at level α according to Theorem 2.2.1 with power function $\hat{M}(\theta)$. If $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ with $h > 0$ then $(N, \hat{\delta})$ is also an MP-test for

$$H_0: \theta = \theta' \quad \text{against} \quad H_1: \theta = \theta''$$

at level $\hat{M}(\theta')$.

Proof. Let (N, δ) be any other test for hypothesis $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ with power function $M(\theta)$, where $M(\theta') \leq \hat{M}(\theta')$. Then, for $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ with $h > 0$ we obtain, cf. (2.128),

$$E_{\theta'}(\hat{\delta} - \delta)(L^h_{N,\theta_0,\theta_1} - c^h_\alpha)\chi_{\{N<\infty\}}$$

$$= E_{\theta''}(\hat{\delta} - \delta)\chi_{\{N<\infty\}} - c^h_\alpha E_{\theta'}(\hat{\delta} - \delta)\chi_{\{N<\infty\}} \geq 0.$$

Hence, $M(\theta') \leq \hat{M}(\theta')$ implies $M(\theta'') \leq \hat{M}(\theta'')$ so that $(N, \hat{\delta})$ is an MP-test for $H_0: \theta = \theta'$ against $H_1: \theta = \theta''$ at level $\hat{M}(\theta')$. ■

Based on this corollary we additionally obtain the following property of an MP-test, which can be interpreted as a certain locally most powerful property.

<u>Lemma 2.2.1</u> Let $(N, \hat{\delta})$ be an MP-test according to Theorem 2.2.1 with power function $\hat{M}(\theta)$. Let (N, δ) be any other test with power function $M(\theta)$ where

$$\hat{M}(\theta^*) = M(\theta^*) \tag{2.131}$$

for any given θ^*, $\underline{\theta} < \theta^* < \bar{\theta}$. Suppose that $\hat{M}$ and M are differentiable at $\theta = \theta^*$. If every ε-neighbourhood of θ^* contains a pair $\theta', \theta'' \in (\underline{\theta}, \bar{\theta})$, $\theta' < \theta^* < \theta''$, so that $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ with $h > 0$, then we have

$$\left.\frac{d\hat{M}(\theta)}{d\theta}\right|_{\theta=\theta^*} \geq \left.\frac{dM(\theta)}{d\theta}\right|_{\theta=\theta^*}. \tag{2.132}$$

Proof. We assume $\left.\frac{d\hat{M}(\theta)}{d\theta}\right|_{\theta=\theta^*} < \left.\frac{dM(\theta)}{d\theta}\right|_{\theta=\theta^*}$. (2.133)

Since (2.131) and $\hat{M}$ and M are differentiable at $\theta = \theta^*$ for each $\varepsilon > 0$ an ε-neighbourhood $U_\varepsilon(\theta^*)$ of θ^* exists so that

$$M(\theta) < \hat{M}(\theta) \quad \text{for} \quad \theta \in U_\varepsilon(\theta^*) \quad \text{and} \quad \theta < \theta^* \tag{2.134}$$

and

$$M(\theta) > \hat{M}(\theta) \quad \text{for} \quad \theta \in U_\varepsilon(\theta^*) \quad \text{and} \quad \theta^* < \theta. \tag{2.135}$$

By assumption there exists a pair $\theta', \theta'' \in U_\varepsilon(\theta^*)$ with $\theta' < \theta^* < \theta''$, $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$ and $h > 0$. Then (2.134) and (2.135) imply

$$M(\theta') < \hat{M}(\theta')$$

and

$$M(\theta'') > \hat{M}(\theta''). \tag{2.136}$$

Otherwise, because of $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$ with $h > 0$ we may apply Corollary 2.2.2. According to this corollary test $(N, \hat{\delta})$ is also an MP-test for $H_o: \theta = \theta'$ against $H_1: \theta = \theta''$ at level $\hat{M}(\theta')$. This implies $M(\theta'') \leq \hat{M}(\theta'')$, which contradicts (2.136). Hence, assumption (2.133) is false and we have (2.132).■

This lemma describes a further optimality property of the MP-test $(N, \hat{\delta})$. Within the class of all tests at sample size N satisfying Lemma 2.2.1 there does not exist a test whose power function at $\theta = \theta^*$ possesses a larger slope than the power function of the MP-test $(N, \hat{\delta})$. This property may be of importance if instead of θ_o and θ_1 the separating-parameter θ^* is given and we are interested in a test for hypothesis

$$H_o: \theta \leq \theta^* \quad \text{against} \quad H_1: \theta > \theta^*.$$

In this context we again refer to Sections 2.1.3 and 2.1.4
Conditions on the differentiability of the power function at $\theta = \theta^*$ have been considered by ABRAHAM [1] and BERK [13] for tests (N, δ) based on a sequence $\{X_n\}_{n \in \Gamma^+}$ of i.i.d. random variables having a density $f_\theta(x)$ w.r.t. some measure μ. These differentiability conditions are fulfilled for the subsequent example.

<u>Example 2.2.3</u>. Let (N, δ) be the MP-test considered in Example 2.2.2. We will show that every ε-neighbourhood of θ^* contains a pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$ satisfying the assumptions of Lemma 2.2.1. In doing this we consider function $\zeta(\theta, \theta^*)$ introduced by Lemma 1.6.4. To given $\varepsilon > 0$ let ζ_ε^* be defined by

$$\zeta_\varepsilon^* = \begin{cases} \min\{\zeta(\theta^* - \varepsilon, \theta^*), \zeta(\theta^* + \varepsilon, \theta^*)\}, & \text{if } \underline{\theta} < \theta^* - \varepsilon \text{ or } \theta^* + \varepsilon < \overline{\theta}, \\ \infty & \text{, otherwise.} \end{cases}$$

Then, the monotonicity properties of $\zeta(\theta, \theta^*)$ in θ, $\zeta(\theta, \theta^*) \geq 0$ for $\theta \in (\underline{\theta}, \overline{\theta})$ and $\zeta(\theta, \theta^*)$ possesses a uniquely determined minimum at $\theta = \theta^*$ imply $\zeta_\varepsilon^* > 0$. Hence, for every ζ_o, $0 < \zeta_o < \zeta_\varepsilon^*$ there exists

a pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$ with

$$\zeta(\theta', \theta^*) = \zeta(\theta'', \theta^*) > 0 \quad \text{and} \quad \theta' < \theta^* < \theta''.$$

Applying Lemma 1.6.5 we have $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ for this pair, where $h > 0$ is determined by (1.65). As shown by [1] and [13], the regularity conditions ensuring the differentiability of the power function at $\theta = \theta^*$ are fulfilled for the distribution family considered here. Thus, Lemma 2.2.1 can be applied and the power function of test $(N, \hat{\delta})$ possesses the largest slope at $\theta = \theta^*$ among all power functions of tests satisfying Lemma 2.2.1. ■

2.3 Unbiased tests

If (N, δ) is a test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ then it is reasonable to restrict our attention only to tests which accept $H_1: \theta = \theta_1$ or $H_0: \theta = \theta_0$, respectively, at least as so often as the corresponding hypothesis is true rather than it is false. In terms of the power function or the OC-function, respectively, that means that for the tests in consideration

$$M(\theta_0) \leq M(\theta_1) \tag{2.137}$$

and

$$Q(\theta_0) \geq Q(\theta_1) \tag{2.138}$$

should be fulfilled.

D e f i n i t i o n 2.3.1. We shall say, the power function or the OC-function of a test (N, δ) for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ satisfies the unbiasedness criterion iff (2.137) or (2.138), respectively, holds. If (2.137) and (2.138) hold simultaneously, the (N, δ) is said to be an unbiased test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$.

The following lemma presents two inequalities concerning the OC- and power function of an MP-test according to Theorem 2.2.1. These inequalities can be used to obtain assertions on the unbiasedness of MP-tests.

L e m m a 2.3.1. Let $(N, \hat{\delta})$ be an MP-test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$ according to Theorem 2.2.1 with power function $M(\theta)$ and OC-function $Q(\theta)$. If $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ with $h > 0$, then

$$M(\theta') \leq c^{-h} M(\theta'') \tag{2.139}$$

and

$$Q(\theta'') \leq c^{h} Q(\theta'). \tag{2.140}$$

The first strict inequality holds if $P_{\theta'}(L_{N,\theta_0,\theta_1} > c_\alpha,\ N < \infty) > 0$, the second one if $P_{\theta'}(L_{N,\theta_0,\theta_1} < c_\alpha,\ N < \infty) > 0$.

P r o o f. For an MP-test $(N, \hat{\delta})$ according to Theorem 2.2.1 we have $L_{N,\theta_0,\theta_1} \leq c_\alpha$ on $\{H_0$ is accepted$\}$ and $L_{N,\theta_0,\theta_1} \geq c_\alpha$ on $\{H_1$ is accepted$\}$. Hence, this lemma is a conclusion of Theorem 2.1.1. ■

We discuss some consequences of this lemma.

(i) Conjugacy and unbiasedness: Let $(N, \hat{\delta})$ be an MP-test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ according to Theorem 2.2.1, where $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$. Then Corollary 2.2.2 implies that this test is also an MP-test for H_0: $\theta = \theta'$ against H_1: $\theta = \theta"$ at level $M(\theta')$. If $c_\alpha \geq 1$, then (2.139) implies

$$M(\theta') \leq M(\theta"),$$

and the power function of $(N, \hat{\delta})$ satisfies the unbiasedness criterion for $\theta_0 = \theta'$ and $\theta_1 = \theta"$. If $c_\alpha \leq 1$, then (2.140) implies

$$Q(\theta') \geq Q(\theta"),$$

and the OC-function of $(N, \hat{\delta})$ satisfies the unbiasedness criterion for $\theta_0 = \theta'$ and $\theta_1 = \theta"$. Putting this together, then $c_\alpha = 1$ and $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$ imply that test $(N, \hat{\delta})$ is an unbiased MP-test for H_0: $\theta = \theta'$ against H_1: $\theta = \theta"$. Especially, because of $(\theta_0,\theta_1) \overset{1}{\sim} (\theta_0,\theta_1)$ every MP-test $(N, \hat{\delta})$ for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ with $c_\alpha = 1$ is unbiased.

(ii) Bounds for c_α: If

$$Q(\theta') > 0 \quad \text{and} \quad M(\theta') > 0, \tag{2.141}$$

then it follows from (2.139) and (2.140)

$$Q(\theta")/Q(\theta') \leq c^h \leq M(\theta")/M(\theta') \tag{2.142}$$

for $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$. Particularly, for $\theta' = \theta_0$ and $\theta" = \theta_1$ we obtain

$$Q(\theta_1)/Q(\theta_0) \leq c_\alpha \leq M(\theta_1)/M(\theta_0). \tag{2.143}$$

If additionally $M(\theta_0) = \alpha$ and $M(\theta_1) = 1 - \beta$ then we have

$$\beta/(1 - \alpha) \leq c_\alpha \leq (1 - \beta)/\alpha . \tag{2.144}$$

(iii) Closedness and unbiasedness: Let $(N, \hat{\delta})$ be a closed MP-test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ where (2.141) and $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ with $h > 0$ holds. Then, instead of (2.142), we obtain

$$(1 - M(\theta"))/(1 - M(\theta')) \leq c_\alpha^h \leq M(\theta")/M(\theta').$$

That implies

$$M(\theta') \leq M(\theta").$$

Since $M(\theta) + Q(\theta) = 1$ for a closed test, we also have

$$Q(\theta') \geq Q(\theta")$$

so that test $(N,\hat{\delta})$ is an unbiased MP-test for H_o: $\theta = \theta'$ against H_1: $\theta = \theta''$ at level $M(\theta')$.

(iv) The unbiasedness of LRTs: Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ be a LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$. Then δ is defined by (2.2). If we may suppose that

$$\sup_{n\in\Gamma^+} B_n < 1 < \inf_{n\in\Gamma^+} A_n \tag{2.145}$$

then terminal decision rule δ can also be written as

$$\delta = \chi_{\{L_{N,\theta_o,\theta_1} \geq 1,\ N<\infty\}}.$$

Hence, if (2.145) holds, then by Theorem 2.2.1 the above LRT is an MP-test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at level $\alpha = E_{\theta_o}\delta\chi_{\{N<\infty\}}$ with $c_\alpha = 1$ and $\gamma_\alpha = 1$. This implies, as already stated in (i), that this test is also an unbiased test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$. Therefore, every WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ with $0<B<1<A<\infty$ is an unbiased MP-LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at level $\alpha = E_{\theta_o}\delta\chi_{\{N<\infty\}}$ with $c_\alpha = 1$ and $\gamma_\alpha = 1$. For the hypotheses H_o: $\theta = \theta'$ against H_1: $\theta = \theta''$ an analogous assertion can be obtained if $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h>0$. ■

Moreover, the following assertion holds for LRTs.

<u>Theorem 2.3.1</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ be an LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$. Then, for every $n\in\Gamma^+$, we have

$$Q_n(\theta_o) \geq Q_n(\theta_1) \quad \text{and} \quad M_n(\theta_o) \leq M_n(\theta_1), \tag{2.146}$$

where

$$Q_n(\theta) = E_\theta\chi_{\{L_{N,\theta_o,\theta_1} \leq B_N,\, N\leq n\}}$$

$$M_n(\theta) = E_\theta\chi_{\{L_{N,\theta_o,\theta_1} \geq A_N,\, N\leq n\}}$$

In particular, (N,δ) is an unbiased LRT.

Proof. The assertion of this theorem can be obtained by mathematical induction. For the sake of abbreviation, we shall write L_n instead of L_{n,θ_o,θ_1} for $n\in\Gamma^+$ here.

(i) We verify (2.146) for $n = 1$: Since $(\theta_o,\theta_1) \overset{h}{\sim} (\theta_o,\theta_1)$ we obtain

$$\begin{aligned} M_1(\theta_o) &= E_{\theta_o}\chi_{\{L_1\geq A_1,\, N=1\}} \\ &\leq A_1^{-1}E_{\theta_o}L_1\chi_{\{L_1\geq A_1,\, N=1\}} \end{aligned}$$

$$= A_1^{-1}E_{\theta_1}\chi\{L_1 \geq A_1, N=1\}$$

$$= A_1^{-1}M_1(\theta_1)$$

by Lemma 1.6.1 for $A_1 \geq 1$. This implies $M_1(\theta_o) \leq M_1(\theta_1)$ for $A_1 \geq 1$. If $A_1 < 1$ holds, we consider

$$\bar{Q}_1(\theta_o) = E_{\theta_o}\chi\{L_1 < A_1, N=1\}$$

$$\geq A_1^{-1}E_{\theta_o}L_1\chi\{L_1 < A_1, N=1\}$$

$$= A_1^{-1}E_{\theta_1}\chi\{L_1 < A_1, N=1\}$$

$$= A_1^{-1}\bar{Q}_1(\theta_1)$$

and we obtain

$$\bar{Q}_1(\theta_o) \geq Q_1(\bar{\theta}_1) \quad \text{for} \quad A_1 < 1. \tag{2.147}$$

By definition of $\bar{Q}_1(\theta)$ we have $\bar{Q}_1(\theta) + M_1(\theta) = 1$ for $\theta \in \Theta$. Therefore, (2.147) implies $M_1(\theta_o) \leq M_1(\theta_1)$ also for $A_1 < 1$. Hence, we obtain $M_1(\theta_o) \leq M_1(\theta_1)$. In an analogous manner we can verify $Q_1(\theta_o) \geq Q_1(\theta_1)$.

(ii) We suppose (2.146) to be true for any $n \in \Gamma^+$: Then, for $A_{n+1} \geq 1$ we have

$$M_{n+1}(\theta_o) = M_n(\theta_o) + E_{\theta_o}\chi\{L_{n+1} \geq A_{n+1}, N=n+1\}$$

$$\leq M_n(\theta_1) + A_{n+1}^{-1}E_{\theta_o}L_{n+1}\chi\{L_{n+1} \geq A_{n+1}, N=n+1\}$$

$$= M_n(\theta_1) + A_{n+1}^{-1}E_{\theta_1}\chi\{L_{n+1} \geq A_{n+1}, N=n+1\}$$

$$\leq M_n(\theta_1) + E_{\theta_1}\chi\{L_{n+1} \geq A_{n+1}, N=n+1\}$$

$$= M_{n+1}(\theta_1).$$

If $A_{n+1} < 1$ holds, we consider

$$\bar{Q}_{n+1}(\theta_o) = Q_n(\theta_o) + E_{\theta_o}\chi\{L_{n+1} < A_{n+1}, N=n+1\}$$

$$\geq Q_n(\theta_1) + A_{n+1}^{-1}E_{\theta_1}\chi\{L_{n+1} < A_{n+1}, N=n+1\}$$

$$= \bar{Q}_{n+1}(\theta_1). \tag{2.148}$$

We again have $\bar{Q}_{n+1}(\theta) + M_{n+1}(\theta) = 1$ for $\theta \in \Theta$. From this by means of (2.148) we obtain $M_{n+1}(\theta_o) \leq M_{n+1}(\theta_1)$ also for $A_{n+1} < 1$, and in a similar manner we can show that

$$Q_{n+1}(\theta_o) \geq Q_{n+1}(\theta_1).$$

Thus (2.146) holds also for n+1. Finally, because of (2.146),

$$M(\theta) = E_\theta \delta \chi_{\{N<\infty\}} = \lim_{n\to\infty} E_\theta \delta \chi_{\{N \leq n\}} = \lim_{n\to\infty} M_n(\theta)$$

and

$$Q(\theta) = E_\theta (1-\delta) \chi_{\{N<\infty\}} = \lim_{n\to\infty} Q_n(\theta)$$

we obtain $M(\theta_o) \leq M(\theta_1)$ and $Q(\theta_o) \geq Q(\theta_1)$ so that (N,δ) is unbiased.■

We notice that in case of $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h>0$ test (N,δ) considered in this theorem is also a unbiased test for H_o: $\theta = \theta'$ against H_1: $\theta = \theta''$.

2.4 Admissible tests

Let (N,δ) be a test for H_o: $\theta \in \Theta_o$ against H_1: $\theta \in \Theta_1$, $\Theta_o, \Theta_1 \subseteq \Theta$, $\Theta_o \cap \Theta_1 = \emptyset$. Then a desirable property of this test is that its power function $M(\theta)$, $\theta \in \Theta$, should be high for values of $\theta \in \Theta_1$ and low for values of $\theta \in \Theta_o$. In order to specify this requirement, we consider so-called admissible tests.

D e f i n i t i o n 2.4.1. A test (N,δ) for H_o: $\theta \in \Theta_o$ against H_1: $\theta \in \Theta_1$, $\Theta_o, \Theta_1 \subseteq \Theta$, $\Theta_o \cap \Theta_1 = \emptyset$, is said to be an admissible test at size (α,β) if to any given α and β, $0<\alpha,\beta<1$, $\alpha+\beta<1$,

$$M(\theta) \leq \alpha \quad \text{for} \quad \theta \in \Theta_o \tag{2.149}$$

and

$$M(\theta) \geq 1-\beta \quad \text{for} \quad \theta \in \Theta_1. \tag{2.150}$$

Assertions concerning the admissibility of a test (N,δ) require certain structural assumptions. In the sequel we investigate how to choose the stopping bounds B and A of WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ to obtain an admissible test at size (α,β). Based on two-sided bounds for the conditional expectation values introduced by Theorem 2.1.1 we shall obtain two-sided bounds for the true risks of our WLRT. These bounds can be used to obtain values for the stopping bounds B and A such that to given α and β test $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ is an admissible test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$.

L e m m a 2.4.1. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT. If $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h>0$,

$$E_{\theta'}(L^h_{N,\theta_o,\theta_1} \chi_{\{N<\infty\}} \mid H_o \text{ is accepted}) \geq B^h \gamma_o(\theta',\theta'') \tag{2.151}$$

and

$$E_{\theta'}(L^h_{N,\theta_0,\theta_1}\chi_{\{N<\infty\}} \mid H_1 \text{ is accepted }) \leq A^h \vartheta_1(\theta',\theta''), \quad (2.152)$$

then the true risks $\alpha(\theta')$ and $\beta(\theta'')$ satisfy

$$\frac{1 - B^h}{\vartheta_1(\theta',\theta'')A^h - B^h} \leq \alpha(\theta') \leq \frac{1 - \vartheta_0(\theta',\theta'')B^h}{A^h - \vartheta_0(\theta',\theta'')B^h} \qquad (2.153)$$

and

$$\frac{\vartheta_0(\theta',\theta'')B^h(A^h - 1)}{A^h - \vartheta_0(\theta',\theta'')B^h} \leq \beta(\theta'') \leq \frac{B^h(\vartheta_1(\theta',\theta'')A^h - 1)}{\vartheta_1(\theta',\theta'')A^h - B^h}. \qquad (2.154)$$

P r o o f. For the sake of abbreviation, we shall write ϑ_0 and ϑ_1 instead of $\vartheta_0(\theta',\theta'')$ and $\vartheta_1(\theta',\theta'')$, respectively. Since (N,δ) is closed, by Theorem 2.1.1, the definition of the true risks $\alpha(\theta')$ and $\beta(\theta'')$, (2.151) and (2.152) we obtain

$$\vartheta_0 B^h \leq \beta(\theta'')/(1 - \alpha(\theta')) \leq B^h \qquad (2.155)$$

and

$$A^h \leq (1 - \beta(\theta''))/\alpha(\theta') \leq \vartheta_1 A^h, \qquad (2.156)$$

where $0 \leq \vartheta_0 \leq 1 \leq \vartheta_1$. From (2.155) we obtain the inequalities

$$\alpha(\theta') + \frac{\beta(\theta'')}{\vartheta_0 B^h} \geq 1 \qquad (2.157)$$

and

$$\alpha(\theta') + \frac{\beta(\theta'')}{B^h} \leq 1. \qquad (2.158)$$

From (2.156) we obtain

$$\frac{\alpha(\theta')}{1/A^h} + \beta(\theta'') \leq 1 \qquad (2.159)$$

and

$$\frac{\alpha(\theta')}{1/\vartheta_1 A^h} + \beta(\theta'') \geq 1. \qquad (2.160)$$

Fig. 4.2.1 illustrates inequalities (2.157) to (2.160). If a point $(\alpha(\theta'),\beta(\theta''))$ satisfies these inequalities, it must belong to the shaded quatriliteral given by points S_1, S_2, S_3 and S_4. These points have the following coordinates:

$$S_1 = ((1 - B^h)/(A^h - B^h),\ B^h(A^h - 1)/(A^h - B^h))$$

$$S_2 = ((1 - \vartheta_0 B^h)/(A^h - \vartheta_0 B^h),\ (\vartheta_0 B^h(A^h - 1)/(A^h - \vartheta_0 B^h))$$

$$S_3 = ((1 - \vartheta_0 B^h)/(\vartheta_1 A^h - \vartheta_0 B^h),\ \vartheta_0 B^h(\vartheta_1 A^h - 1)/(\vartheta_1 A^h - \vartheta_0 B^h))$$

$$S_4 = ((1 - B^h)/(\vartheta_1 A^h - B^h),\ B^h(\vartheta_1 A^h - 1)/(\vartheta_1 A^h - B^h)).$$

Hence, risk $\alpha(\theta')$ ranges between the abscissa of S_4 and the abscissa of S_2, risk $\beta(\theta'')$ ranges between the ordinates of S_2 and S_4. ■

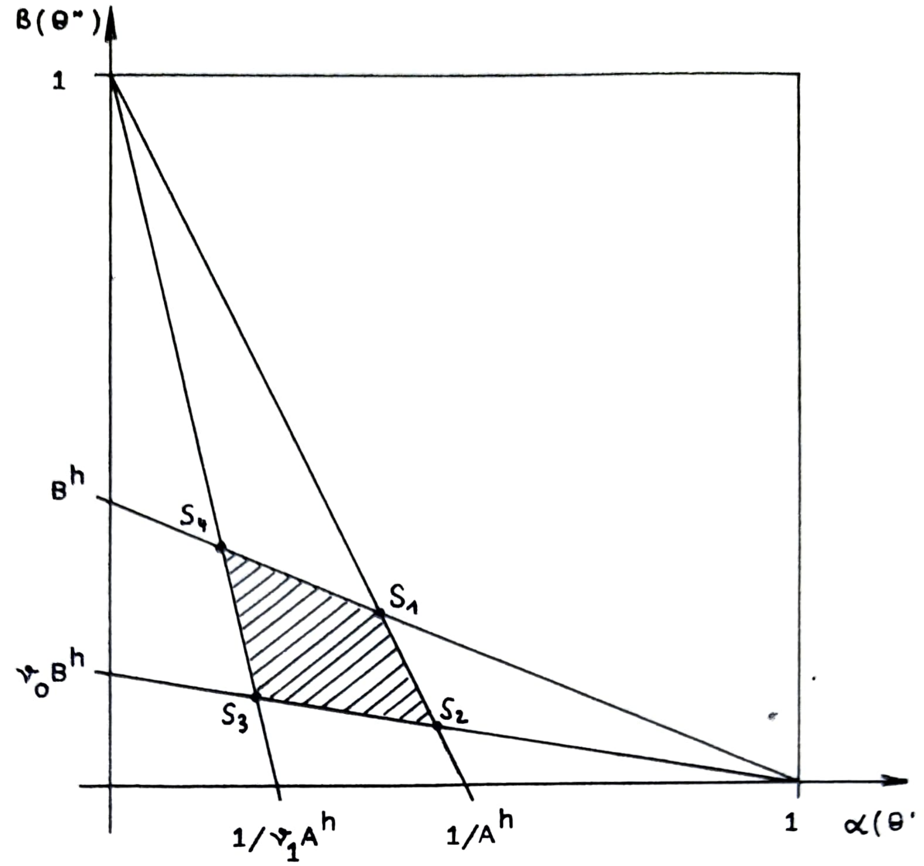

Fig. 4.2.1 Graphical representation of inequalities (2.167) to (2.160)

By means of inequalities (2.153) and (2.154) we obtain the following admissibility criterion for WLRTs.

L e m m a 2.4.2. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a closed WLRT where

$$E_{\theta_0}(L_{N,\theta_0,\theta_1}\chi\{N<\infty\} \mid H_0 \text{ is accepted}) \geqslant B\nu_0(\theta_0,\theta_1) \quad (2.161)$$

and

$$E_{\theta_0}(L_{N,\theta_0,\theta_1}\chi\{N<\infty\} \mid H_1 \text{ is accepted}) \leqslant A\nu_1(\theta_0,\theta_1). \quad (2.162)$$

Then (N,δ) is an admissible test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ at size (α,β) if the stopping bounds B and A satisfy

$$A \geqslant \frac{1}{\alpha}\,(1 - (1-\alpha)\nu_0(\theta_0,\theta_1)\cdot B) \quad (2.163)$$

and

$$B \leqslant \beta\,\Big(1 + \frac{1-\beta}{\nu_1(\theta_0,\theta_1)\cdot A + \beta - 1}\Big). \quad (2.164)$$

P r o o f. Since $(\theta_0,\theta_1) \overset{1}{\sim} (\theta_0,\theta_1)$ it follows from (2.153) and (2.154) that test $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ is an admissible test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ at size (α,β) if

$$\alpha(\theta_0) \leq (1 - \vartheta_0(\theta_0,\theta_1)B)/(A - \vartheta_0(\theta_0,\theta_1)B) \leq \alpha \tag{2.165}$$

and

$$\beta(\theta_1) \leq B(\vartheta_1(\theta_0,\theta_1)A - 1)/(\vartheta_1(\theta_0,\theta_1)A - B) \leq \beta . \tag{2.165'}$$

The right-hand sides of these inequalities immediately provide (2.163) and (2.164).■

Since $0 \leq \vartheta_0(\theta_0,\theta_1) \leq 1 \leq \vartheta_1(\theta_0,\theta_1)$ it follows from (2.163) and (2.164) that under the conditions of this lemma WLRT

$$(N, \delta) = \left\{L_{n,\theta_0,\theta_1}, \beta, \frac{1}{\alpha}\right\}_{n \in \Gamma^+} \tag{2.166}$$

is always an admissible test for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ at size (α, β). More precisely, the following assertion holds.

<u>Corollary 2.4.1</u>. Under the conditions of Lemma 2.4.2 the smallest value A_0 for A and the largest value B_0 for B so that (2.162) and (2.163) hold are given by

$$A_0 = \frac{c_1 + c_3 - c_2\beta}{2} + \sqrt{\frac{(c_1 + c_3 - c_2\beta)^2}{4} - c_1c_3} \tag{2.167}$$

and

$$B_0 = \beta \frac{A_0}{A_0 - c_3}, \tag{2.168}$$

where c_1, c_2 and c_3 are determined by

$$c_1 = \frac{1}{\alpha}, \quad c_2 = \frac{1-\alpha}{\alpha}\vartheta_0(\theta_0,\theta_1) \quad \text{and} \quad c_3 = \frac{1-\beta}{\vartheta_1(\theta_0,\theta_1)}. \tag{2.169}$$

Proof. As shown in the proof of Lemma 2.4.2, the considered test is admissible at size (α, β) if (2.165) and (2.165') hold. From (2.165) follows

$$\frac{A}{1/\alpha} + \frac{B}{1/(1-\alpha)\vartheta_0} \geq 1, \tag{2.170}$$

from (2.165') follows

$$B \leq \beta \frac{A}{A - (1-\beta)/\vartheta_1}. \tag{2.171}$$

In an (A,B)-coordinate system equation (2.171) describes a hyperbola whose asymptotes are given by the equations

$$A = (1 - \beta)/\vartheta_1 \quad \text{and} \quad B = \beta .$$

Since $0 \leq \vartheta_0 \leq 1 \leq \vartheta_1$ and $0 < \alpha, \beta < 1$ it can be shown that the straight line given by (2.170) and the hyperbola given by (2.171) have one and only one common point for $0 < B < 1$ and $1 < A < \infty$. This is illustrated

in Fig. 4.2.2. Hence, test $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ is admissible if point (B,A) satisfies inequalities (2.170) and (2.171), and this is fulfilled, if point (B,A) belongs to the shaded region of Fig. 4.2.2. By (2.169) we obtain

and
$$A = c_1 - c_2 B$$
$$B = \beta A/(A - c_3) \tag{2.172}$$

for the equations (2.170) and (1.171). This provides the quadratic equation

$$A^2 + (c_2\beta - c_1 - c_3)A + c_1c_3 = 0.$$

For $1 < A < \infty$ we obtain solution (2.167), and, together with (2.172), we obtain (2.168). ■

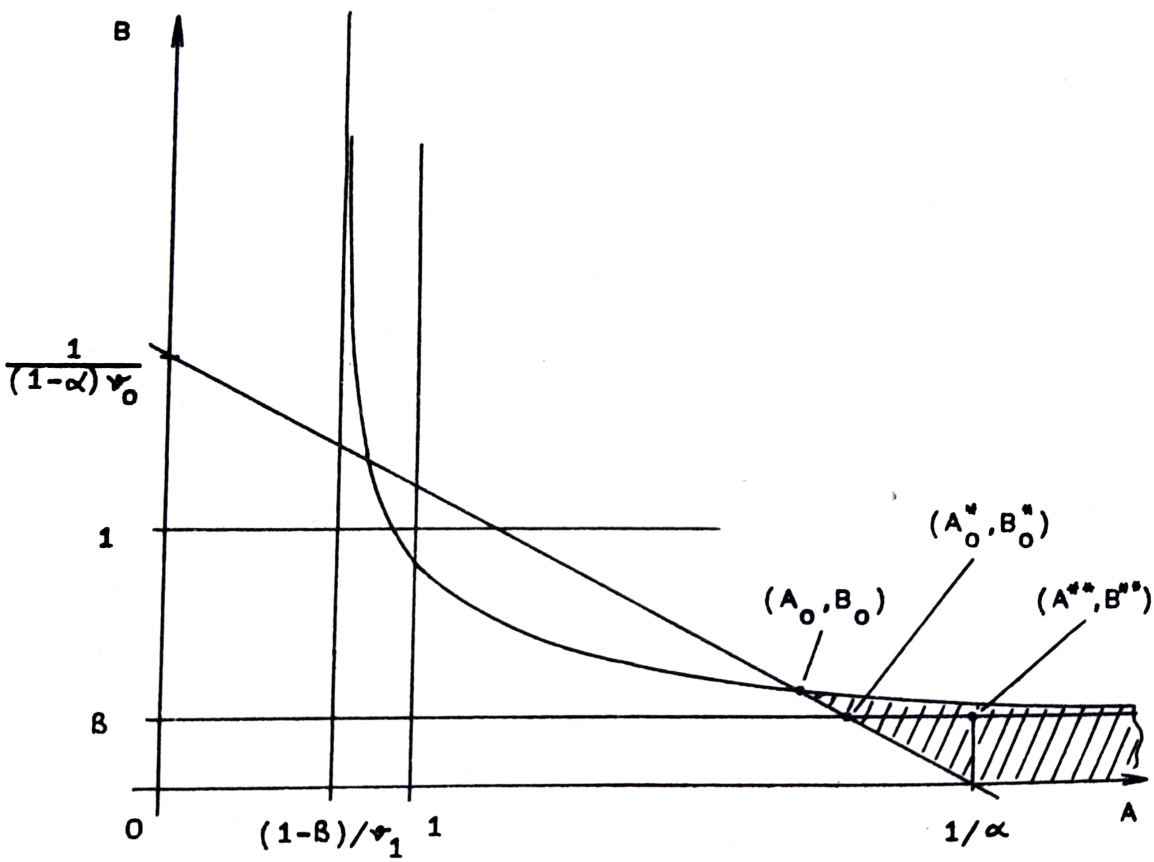

Fig. 4.2.2. Graphical representation of inequalities (2.170) and (2.171)

As already stated above, test (2.166) is an admissible test. This also follows from the above corollary if we put $\vartheta_0(\theta_0,\theta_1) = 0$ and $\vartheta_1(\theta_0,\theta_1) = +\infty$.

For some cases - compare the examples at the end of this section -

it will be easier to compute $\vartheta_0(\theta_0,\theta_1)$ than $\vartheta_1(\theta_0,\theta_1)$. If in such case's we put $\vartheta_1(\theta_0,\theta_1) = \infty$, then we obtain

$$A_0^* = c_1 - c_2\beta = \frac{1}{\alpha} - \frac{(1-\alpha)\beta\vartheta_0(\theta_0,\theta_1)}{\alpha} \tag{2.173}$$

and

$$B_0^* = \beta \tag{2.174}$$

by (2.167) and (2.168). If we choose the stopping bounds in such a manner, then this may already be an essential improvement in comparison with $A = A^{**} = 1/\alpha$ and $B = B^{**} = \beta$.

Now we consider a method for determining quantities $\vartheta_0(\theta',\theta")$ and $\vartheta_1(\theta',\theta")$ introduced by Lemma 2.4.1.

<u>Lemma 2.4.3</u>. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ be a closed WLRT with

$$\ln L_{n,\theta_0,\theta_1} = \sum_{i=1}^{n} Y_i, \quad n\in\Gamma^+,$$

where $\{Y_n\}_{n\in\Gamma^+}$ are i.i.d. random variables. If $(\theta',\theta") \overset{h}{\sim} (\theta_0,\theta_1)$ with $h>0$, then (2.151) and (2.152) holds for

$$\vartheta_0(\theta',\theta") = \inf_{0<\xi\leq 1} \frac{1}{\xi}\cdot\frac{P_{\theta"}(L_{1,\theta',\theta"} \leq \xi)}{P_{\theta'}(L_{1,\theta',\theta"} \leq \xi)} \tag{2.175}$$

and

$$\vartheta_1(\theta',\theta") = \sup_{1\leq\xi<\infty} \frac{1}{\xi}\,\frac{P_{\theta"}(L_{1,\theta',\theta"} \geq \xi)}{P_{\theta'}(L_{1,\theta',\theta"} \geq \xi)}\ . \tag{2.176}$$

P r o o f. Let Z_N be defined by $Z_N = \ln L_{N,\theta_0,\theta_1}$. Then we obtain

$$\underline{B} = E_{\theta'}(L^h_{N,\theta_0,\theta_1}\chi\{N<\infty\} \mid H_0 \text{ is accepted })$$

$$= E_{\theta'}(\exp(hZ_N) \mid Z_N \leq \ln B)$$

$$= E_{\theta'}(\exp(hZ_{N-1} + hY_N) \mid Y_N \leq \ln B - Z_{N-1}).$$

Denote by $F_{Z_{N-1}}(z)$ the distribution function of Z_{N-1} for θ', then because of $\ln B < Z_{N-1} < \ln A$ we obtain

$$\underline{B} = \int_b^a E_{\theta'}(\exp(hZ_{N-1} + hY_N \mid Y_N \leq \ln B - Z_{N-1}, Z_{N-1}=z)dF_{Z_{N-1}}(z)$$

$$= \int_b^a \exp(hz)E_{\theta'}(\exp(hY_N \mid Y_N \leq \ln B - z, Z_{N-1}=z)dF_{Z_{N-1}}(z).$$

The $\{Y_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables. This implies

$$E_{\theta'}(\exp(hY_N) \mid Y_N \leq \ln B - z, Z_{N-1} = z)$$

$$= \sum_{n\in\Gamma^+} E_{\theta'}(\exp(hY_N) \mid Y_N \leq \ln B - z, Z_{N-1}=z, N=n) P_{\theta'}(N = n)$$

$$= \sum_{n\in\Gamma^+} E_{\theta'}(\exp(hY_n) \mid Y_n \leq \ln B - z, Z_{n-1}=z, N=n) P_{\theta'}(N=n)$$

$$= \sum_{n\in\Gamma^+} E_{\theta'}(\exp(hY_n) \mid Y_n \leq \ln B - z, N=n) P_{\theta'}(N = n)$$

$$= \sum_{n\in\Gamma^+} E_{\theta'}(\exp(hY_1) \mid Y_1 \leq \ln B - z, N=n) P_{\theta'}(N = n)$$

$$= E_{\theta'}(\exp(hY_1) \mid Y_1 \leq \ln B - z),$$

and by $h > 0$ we obtain

$$\underline{B} = \int_b^a \exp(hz) E_{\theta'}(\exp(hY_1) \mid Y_1 \leq \ln B - z) dF_{Z_{N-1}}(z)$$

$$= \int_b^a \exp(hz) E_{\theta'}(\exp(hY_1) \mid \exp(hY_1) \leq \exp(h(\ln B - z))) dF_{Z_{N-1}}(z).$$

We substitute

$$\xi = \exp(h(\ln B - z)) \quad \text{for} \quad 0 < z < \infty.$$

Then $\ln B < z < \ln A$ is equivalent to $1 > \xi > \exp(h(\ln B - \ln A)) > 0$, and we obtain

$$\underline{B} = \int_{\xi=0}^{1} \exp(hb)\xi^{-1} E_{\theta'}(\exp(hY_1) \mid \exp(hY_1) \leq \xi) dF_{Z_{N-1}}\left(\ln B - \frac{\ln \xi}{h}\right)$$

$$\geq \exp(hb) \inf_{0<\xi\leq 1} \xi^{-1} E_{\theta'}(\exp(hY_1) \mid \exp(hY_1) \leq \xi) \cdot 1$$

$$= B^h \inf_{0<\xi\leq 1} \xi^{-1} E_{\theta'}(L^h_{1,\theta_0,\theta_1} \mid L^h_{1,\theta_0,\theta_1} \leq \xi)$$

$$= B^h \inf_{0<\xi\leq 1} \xi^{-1} E_{\theta'}(L_{1,\theta',\theta''} \mid L_{1,\theta',\theta''} \leq \xi)$$

$$= B^h \inf_{0<\xi\leq 1} \xi^{-1} \frac{\int_{\{L_{1,\theta',\theta''} \leq \xi\}} L_{1,\theta',\theta''} dP_{\theta'}}{P_{\theta'}(L_{1,\theta',\theta''} \leq \xi)}$$

$$= B^h \inf_{0<\zeta\leq 1} \zeta^{-1} \frac{P_{\theta''}(L_{1,\theta',\theta''}\leq\zeta)}{P_{\theta'}(L_{1,\theta',\theta''}\leq\zeta)},$$

so that (2.151) holds, where $\vartheta_o(\theta',\theta'')$ is determined by (1.175). Analogously, we can show that (2.152) holds if $\vartheta_1(\theta',\theta'')$ is given by (2.176). ■

This lemma, for instance, can be applied if $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ is a WLRT based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density $f_\theta(x)$, $\theta\in\Theta$, $x\in\mathfrak{X}$. Then we have

$$\ln L_{n,\theta_o,\theta_1} = \sum_{i=1}^{n} \ln(f_{\theta_1}(X_i)/f_{\theta_o}(X_i))$$

and

$$Y_n = \ln(f_{\theta_1}(X_n)/f_{\theta_o}(X_n)), \quad n\in\Gamma^+.$$

If, moreover, the distribution of X_1 belongs to a one-parametric exponential family, quantities $\vartheta_o(\theta',\theta'')$ and $\vartheta_1(\theta',\theta'')$ can be determined as follows.

<u>L e m m a 2.4.4</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ be a closed WLRT based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density

$$f_\theta(x) = h(x)\exp(d(\theta)\cdot x - c(\theta)), \quad x\in\mathfrak{X},\ \theta\in\Theta. \tag{2.177}$$

Suppose that d is strictly monotonically increasing in θ on Θ. If $\theta'<\theta''$ and $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h>0$, then we obtain

$$\vartheta_o(\theta',\theta'') = \exp(c(\theta'')-c(\theta')) \inf_{-\infty<\xi\leq\xi^*} \left\{\exp(-(d(\theta'')-d(\theta'))\xi)\frac{P_{\theta''}(X_1\leq\xi)}{P_{\theta'}(X_1\leq\xi)}\right\}, \tag{2.178}$$

$$\vartheta_1(\theta',\theta'') = \exp(c(\theta'')-c(\theta')) \sup_{\xi^*\leq\xi<\infty} \left\{\exp(-(d(\theta'')-d(\theta'))\xi)\frac{P_{\theta''}(X_1\geq\xi)}{P_{\theta'}(X_1\geq\xi)}\right\} \tag{2.179}$$

with

$$\xi^* = (c(\theta_1)-c(\theta_o))/(d(\theta_1)-d(\theta_o)). \tag{2.180}$$

P r o o f. We substitute variable ζ of (1.175) by

$$\zeta = f_{\theta''}(\xi)/f_{\theta'}(\xi) = \exp((d(\theta'')-d(\theta'))\xi - (c(\theta'')-c(\theta'))). \tag{2.181}$$

By $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h>0$, $\theta'<\theta''$, (2.177) and d is strictly monotonically increasing we obtain

$$P_{\theta'}(L_{1,\theta',\theta''}\leq\zeta) = P_{\theta'}(X_1\leq\xi). \tag{2.182}$$

Furthermore, inequality

$$0 < \zeta \leq 1 \tag{2.183}$$

of (2.175) is equivalent to

$$-\infty < \xi < (c(\theta'') - c(\theta'))/(d(\theta'') - d(\theta')).$$

By Lemma 1.6.3, (1.31) and (1.32) this inequality is furthermore equivalent to

$$-\infty < \xi < (c(\theta_1) - c(\theta_0))/(d(\theta_1) - d(\theta_0)) = \xi^*. \tag{2.184}$$

Then (2.178) follows from (2.175), (2.181), (2.182) and the equivalence of (2.183) and (2.184). Analogously, we obtain (2.179). ■

We note, if d is strictly monotonically decreasing in θ on Θ, then, instead of (2.178) and (2.179), we obtain

$$\psi_0(\theta',\theta'') = \exp(c(\theta'')-c(\theta')) \inf_{\xi^* \leq \xi < \infty} \left\{\exp(-(d(\theta'')-d(\theta'))\xi)\frac{P_{\theta''}(X_1 \geq \xi)}{P_{\theta'}(X_1 \geq \xi)}\right\} \tag{2.185}$$

and

$$\psi_1(\theta',\theta'') = \exp(c(\theta'')-c(\theta')) \sup_{-\infty < \xi \leq \xi^*} \left\{\exp(-(d(\theta'')-d(\theta'))\xi)\frac{P_{\theta''}(X_1 \leq \xi)}{P_{\theta'}(X_1 \leq \xi)}\right\}. \tag{2.186}$$

Moreover, we remark that variable ξ used in Lemma 2.4.4 is a continuous variable, even if the random variables $\{X_n\}_{n\in\Gamma^+}$ are discrete. If we have a sequence of integer-valued random variables, then Lemma 2.4.4 can be modified as follows.

C o r o l l a r y 2.4.2. Suppose that Lemma 2.4.3 holds. If $\mathfrak{X} \subseteq \Gamma$, then

$$\psi_0(\theta',\theta'') = \min\left\{\frac{P_{\theta''}(X_1 \leq \xi^*)}{P_{\theta'}(X_1 \leq \xi^*)}, \inf_{x\in\mathfrak{X},\, x<\xi^*}\left\{\frac{f_{\theta'}(x+1)\; P_{\theta''}(X_1 \leq x)}{f_{\theta''}(x+1)\; P_{\theta'}(X_1 \leq x)}\right\}\right\} \tag{2.187}$$

and

$$\psi_1(\theta',\theta'') = \max\left\{\frac{P_{\theta''}(X_1 \geq \xi^*)}{P_{\theta'}(X_1 \geq \xi^*)}, \sup_{x\in\mathfrak{X},\, \xi^*<x}\left\{\frac{f_{\theta'}(x-1)\; P_{\theta''}(X_1 \geq x)}{f_{\theta''}(x-1)\; P_{\theta'}(X_1 \geq x)}\right\}\right\} \tag{2.188}$$

where ξ^* is determined by (2.180).

P r o o f. This corollary immediately follows from Lemma 2.4.3. ■

With respect to possible applications, the following special case may be of particular interest.

C o r o l l a r y 2.4.3. Suppose that Lemma 2.4.4 holds. If $\mathfrak{X} \subseteq \Gamma_0^+$, $0 \in \mathfrak{X}$ and $0 < \xi^* < 1$, then

$$\psi_0(\theta',\theta'') = \exp(-(c(\theta'') - c(\theta'))). \tag{2.189}$$

P r o o f. By $0 < \xi^* < 1$, $0 \in \mathfrak{X}$ and (2.178) we obtain

$$\vartheta_0(\theta',\theta'') = \exp(c(\theta'')-c(\theta')-(d(\theta'')-d(\theta'))\xi^*)\frac{P_{\theta''}(X_1 = 0)}{P_{\theta'}(X_1 = 0)} .$$

By Lemma 1.6.3, (2.180) and (2.177) we obtain

$\exp(c(\theta'')-c(\theta')-(d(\theta'')-d(\theta'))\xi^*) = 1$ and $P_\theta(X_1=0) = h(0)\exp(-c(\theta))$.

This, together with (2.190), implies (2.189).■

E x a m p l e 2.4.1. We consider the computation of $\vartheta_0(\theta',\theta'')$ and $\vartheta_1(\theta',\theta'')$ for some special cases of (2.177). We refer in this context also to Examples 1.6.1 and 2.1.0.

(i) The binomial proportion. We suppose $0<\theta_0<\theta_1<1$, then by (2.180) we obtain

$$\xi^* = \ln\frac{1-\theta_0}{1-\theta_1}\Big/\ln\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)} \quad \text{with } 0<\xi^*<1.$$

Applying Corollary 2.4.3 we obtain

$$\vartheta_0(\theta',\theta'') = \frac{1-\theta''}{1-\theta'} = \left(\frac{1-\theta_1}{1-\theta_0}\right)^h .$$

In order to obtain $\vartheta_1(\theta',\theta'')$, we notice that

$$P_\theta(X_1 \geqslant \xi) \leqslant P_\theta(X_1 \geqslant \xi^*) = P_\theta(X_1 = 1) = \theta \quad \text{for } \xi^* \leqslant \xi .$$

Then, by (2.188) we obtain

$$\vartheta_1(\theta',\theta'') = \theta''/\theta' = (\theta_1/\theta_0)^h .$$

(ii) The Poisson mean. We obtain

$$\xi^* = (\theta_1 - \theta_0)/\ln(\theta_1/\theta_0).$$

If we may suppose that $0<\theta_0<\theta_1<1$, then we have $0<\xi^*<1$. Applying Corollary 2.4.3 we obtain

$$\vartheta_0(\theta',\theta'') = \exp(\theta' - \theta'') = \exp(h(\theta_0 - \theta_1)).$$

By (2.188) we obtain

$$\vartheta_1(\theta',\theta'') = \max\left\{\frac{1-\theta'' e^{-\theta''}}{1-\theta' e^{-\theta'}},\ e^{-(\theta'-\theta'')} \sup_{x\in\mathfrak{X},\,1<x}\left\{\left(\frac{\theta'}{\theta''}\right)^{x-1}\frac{P_{\theta''}(X_1 \geqslant x)}{P_{\theta'}(X_1 \geqslant x)}\right\}\right\} .$$

(iii) The normal mean. We consider WLRT for mean θ with hypotheses

$$H_0:\ \theta = -\theta_1 \quad\text{and}\quad H_1:\ \theta = \theta_1,\quad \theta_1 > 0,$$

and assume that $\sigma^2 = 1$. Then we have $d(\theta) = \theta$ and $c(\theta) = \theta^2/2$. This implies $\xi^* = 0$. We start with $\vartheta_1(\theta',\theta'')$ and remark that here $(\theta',\theta'') \overset{h}{\sim} (-\theta_1,\theta\ )$ with $h>0$ implies $\theta' = -\theta''$ and $\theta'' = h\theta_1$. Then, by (2.179) we obtain

$$\vartheta_1(\theta',\theta'') = \vartheta_1(-\theta'',\theta'') = \sup_{0\leqslant\xi<\infty}\left\{\exp(-2\theta''\xi)\phi(\theta''-\xi)/\phi(-\theta''-\xi)\right\}.$$

Let y_1 be defined by $y_1 = \exp(-2\theta''\xi)$ for $\xi \geq 0$. Then y_1 is a decreasing function in ξ. Let y_2 be defined by

$$y_2 = \phi(\theta'' - \xi)/\phi(-\theta'' - \xi), \qquad \phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-z^2/2)dz.$$

Then we have

$$y_2 = 1 + \int_{-\theta''-\xi}^{\theta''-\xi} \varphi(z)dz \Big/ \phi(-\theta''-\xi), \qquad \varphi(z) = \frac{1}{\sqrt{2\pi}} \exp(-z^2/2),$$

and we obtain

$$\begin{aligned} y_2' &= (-\varphi(\theta''-\xi) + \varphi(-\theta''-\xi)\phi(\theta''-\xi))/(\phi(-\theta''-\xi))^2 \\ &\leq (-\varphi(\theta''-\xi) + \varphi(-\theta''-\xi)\cdot 1)/(\phi(-\theta''-\xi))^2 \\ &\leq 0 \quad \text{for } \xi \geq 0 \end{aligned}$$

for the first derivative of y_2 w.r.t. ξ. Hence, also y_2 is a decreasing function in ξ on $\xi \geq 0$. This implies

$$\gamma_1(\theta',\theta'') = \frac{\phi(\theta'')}{\phi(-\theta'')} = \frac{\phi(h\theta_1)}{\phi(-h\theta_1)}.$$

For $\gamma_o(\theta',\theta'')$ we obtain

$$\gamma_o(\theta'\ \theta'') = \gamma_o(-\theta'',\theta'') = \inf_{-\infty<\xi\leq 0} \left\{\exp(-2\theta''\xi) \frac{\phi(\xi - \theta'')}{\phi(\xi + \theta'')}\right\}.$$

We substitute $\xi = -\eta$ and obtain with $-\theta'' = \theta' < 0$

$$\gamma_o(\theta',\theta'') = \inf_{0\leq\eta<\infty} \left\{\exp(-2\theta'\eta) \frac{\phi(-\eta + \theta')}{\phi(-\eta - \theta')}\right\},$$

where $\exp(-2\theta'\eta)$ and $\phi(-\eta + \theta')/\phi(-\eta - \theta')$ are increasing functions in η. Then we have

$$\gamma_o(\theta',\theta'') = \frac{\phi(\theta')}{\phi(-\theta')} = \frac{\phi(-\theta'')}{\phi(\theta'')} = (\gamma_1(\theta',\theta''))^{-1}.$$

The assumption $\sigma^2 = 1$ and the symmetrical choice of the hypotheses are no restriction. If, for instance, we have a sequence $\{\hat{x}_n\}_{n\in\Gamma^+}$ of independent $N(\hat{\theta},\sigma^2)$-distributed random variables with known variance $\sigma^2 > 0$ and if we want to test hypothesis

$$H_o: \hat{\theta} = \hat{\theta}_o \quad \text{against} \quad H_1: \hat{\theta} = \hat{\theta}_1, \quad \hat{\theta}_o < \hat{\theta}_1, \tag{2.191}$$

then we can always use the following transformations:

$$x_n = \left(\hat{x}_n - \frac{\hat{\theta}_o + \hat{\theta}_1}{2}\right)\Big/\sigma \quad \text{for } n \in \Gamma^+, \tag{2.192}$$

$$\theta = \left(\hat{\theta} - \frac{\hat{\theta}_o + \hat{\theta}_1}{2}\right)\Big/\sigma. \tag{2.193}$$

Then, the random variables $\{x_n\}_{n\in\Gamma^+}$ are independent $N(\theta,1)$-distributed random variables and hypotheses (2.191) are equivalent to

$$H_o: \theta = -\theta_1 = -\frac{\hat{\theta}_1 - \hat{\theta}_o}{2\sigma} \quad \text{and} \quad H_1: \theta = \theta_1 = \frac{\hat{\theta}_1 - \hat{\theta}_o}{2\sigma}, \ \theta_1 > 0, \tag{2.194}$$

which corresponds to our above requirement.

Indeed, we remark that for WLRTs concerning the normal mean the stopping bounds B_o and A_o given by Corollary 2.4.1 should be used only if α and β are different. For the symmetrical case $\alpha = \beta$, we refer to Section 2.6.1, Lemma 2.1.8 and Corollary 2.1.5.

(iv) The exponential distribution. We suppose

$$f_\theta(x) = \theta \exp(-\theta x), \quad x \in (0,\infty), \quad \theta \in (\theta,\infty).$$

Then we obtain

$$\xi^* = (\ln(\theta_1/\theta_o))/(\theta_1 - \theta_o)$$

where we assume that $\theta_o < \theta_1$. We note that $d(\theta) = -\theta$ is a decreasing function in θ on $(0,\infty)$. Hence, $\vartheta_o(\theta',\theta'')$ and $\vartheta_1(\theta',\theta'')$ must be determined by (2.185) and (2.186). Since $P_\theta(X_1 \geq \xi) = \exp(-\theta\xi)$ we obtain

$$\vartheta_o(\theta',\theta'') = (\theta'/\theta'') = (\theta_o/\theta_1)^h.$$

Further, we obtain

$$\vartheta_1(\theta',\theta'') = \frac{\theta'}{\theta''} \sup_{-\infty<\xi\leq\xi^*} \left\{ \frac{\exp(\theta''\xi) - 1}{\exp(\theta'\xi) - 1} \right\}.$$

For $\theta' < \theta''$ function $y = (\exp(\theta''\xi) - 1)/(\exp(\theta'\xi) - 1)$ is a non-decreasing function in ξ on R^1. This provides

$$\vartheta_1(\theta',\theta'') = \frac{\theta'}{\theta''} \frac{\exp(\theta''\xi^*) - 1}{\exp(\theta'\xi^*) - 1}.$$

A numerical example. We choose

$$\theta_o = 1, \quad \theta_1 = 2, \quad \alpha = 0.05 \quad \text{and} \quad \beta = 0.05.$$

and obtain

$$\xi^* = \ln 2 = 0.6931, \ \vartheta_o(\theta_o,\theta_1) = 0.5 \text{ and } \vartheta_1(\theta_o,\theta_1) = 1.5.$$

According to Corollary 2.4.1 we obtain

$$B_o = 0.0517 \quad \text{and} \quad A_o = 19.5091,$$

and test $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B_o, A_o\}_{n\in\Gamma^+}$ is an admissible test for $H_o: \theta = 1$ against $H_1: \theta = 2$ at level $(0.05,0.05)$. In comparison with it we consider the WALD approximations for the stopping bounds and obtain

$$B^* = 0.0526 \quad \text{and} \quad A^* = 19.$$

We remark that these approximations do not satisfy the admissibility criterion given by Lemma 2.4.2.

2.5 Monotone likelihood ratio families

The most powerful property and the unbiasedness of an LRT considered in Sections 2.2 and 2.3 were obtained without any additional structural assumptions concerning the likelihood ratios. As already stated in Example 1.4.1, if (N, δ) is a test based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables with a distribution from an exponential family, then a sequence of statistics $\{T_n\}_{n\in\Gamma^+}$ exists so that L_{n,θ_0,θ_1} is a function of T_n for every $n\in\Gamma^+$. Now we discuss some consequences of such a representation possibility concerning the structure of LRTs and especially of WLRTs. Furtheron, we present a monotonicity criterion for the monotonicity of the power function of an LRT.

<u>D e f i n i t i o n 2.5.1</u>. A one-parameter family $\mathcal{P} = \{P_\theta, \theta \in (\underline{\theta},\bar{\theta})\}$ is said to be a <u>monotone likelihood ratio family</u> (MLRF) if for every $n\in\Gamma^+$ there exists a statistic $T_n : \Omega \to \mathcal{T} \subseteq R^1$ so that, for every pair $\theta', \theta'' \in (\underline{\theta},\bar{\theta})$, $\theta' < \theta''$, a measurable function $g_{n,\theta',\theta''} : \mathcal{T} \to R^1$ exists such that

$$L_{n,\theta',\theta''} = g_{n,\theta',\theta''}(T_n) \tag{2.195}$$

and $g_{n,\theta',\theta''}(t)$ is strictly monotonous in t on $\mathcal{T}$.

If $\mathcal{P}$ is an MLRF, then the sample size and the terminal decision rule of any given LRT $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ can be represented as follows.

<u>L e m m a 2.5.1</u>. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ be an LRT for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, $\underline{\theta} < \theta_0 < \theta_1 < \bar{\theta}$. If $\mathcal{P} = \{P_\theta, \theta\in(\underline{\theta},\bar{\theta})\}$ is an MLRF where g_{n,θ_0,θ_1} is an increasing function for every $n\in\Gamma^+$, then we have

$$N = \begin{cases} \inf\{n \geq 1: T_n \notin (c_{n,\theta_0,\theta_1}, d_{n,\theta_0,\theta_1})\}, & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{2.196}$$

and

$$\delta = \chi_{\{T_N \geq d_{N,\theta_0,\theta_1},\, N<\infty\}}, \tag{2.197}$$

where c_{n,θ_0,θ_1} and d_{n,θ_0,θ_1} are determined by

$$c_{n,\theta_0,\theta_1} = g^{-1}_{n,\theta_0,\theta_1}(B_n) \quad\text{and}\quad d_{n,\theta_0,\theta_1} = g^{-1}_{n,\theta_0,\theta_1}(A_n) \tag{2.198}$$

respectively, $n\in\Gamma^+$.

P r o o f. Since $\mathcal{P}$ is an MLRF, we have $L_{n,\theta_0,\theta_1} = g_{n,\theta_0,\theta_1}(T_n)$, $n\in\Gamma^+$,

The functions g_{n,θ_o,θ_1}, $n\in\Gamma^+$, are assumed to be increasing, therefore inequalities

$$B_n < L_{n,\theta_o,\theta_1} < A_n \quad \text{and} \quad L_{n,\theta_o,\theta_1} \geq A_n, \quad n\in\Gamma^+,$$

are equivalent to

$$c_{n,\theta_o,\theta_1} = g^{-1}_{n,\theta_o,\theta_1}(B_n) < T_n < g^{-1}_{n,\theta_o,\theta_1}(A_n) = d_{n,\theta_o,\theta_1}$$

and

$$T_n \geq g^{-1}_{n,\theta_o,\theta_1}(A_n) = d_{n,\theta_o,\theta_1}, \quad n\in\Gamma^+,$$

respectively. This implies (2.196) and (2.197). ■

A similar assertion can be obtained if g_{n,θ_o,θ_1} is a strictly monotonically decreasing function for every $n\in\Gamma^+$.

<u>Example 2.5.1</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ be an LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$, $\underline{\theta} < \theta_o < \theta_1 < \overline{\theta}$, based on a sequence of i.i.d. random variables $\{X_n\}_{n\in\Gamma^+}$ having density

$$f_\theta(x) = h(x)\cdot\exp(d(\theta)t(x) - c(\theta)).$$

We suppose that $d(\theta)$ is strictly monotonically increasing in θ on $(\underline{\theta},\overline{\theta})$. Then, for every pair $\theta',\theta'' \in (\underline{\theta},\overline{\theta})$ with $\theta' < \theta''$ and $n\in\Gamma^+$ we have

$$L_{n,\theta',\theta''} = \exp((d(\theta'') - d(\theta'))T_n - n(c(\theta'') - c(\theta'))$$

with

$$T_n = \sum_{n=1}^{n} t(X_i).$$

This implies

$$g_{n,\theta',\theta''}(t) = \exp((d(\theta'') - d(\theta'))t - n(c(\theta'') - c(\theta')), \quad (2.198)$$

which is a strictly monotonically increasing function in t on R^1. Hence, $\mathcal{P}$ forms an MLRF. For $g^{-1}_{n,\theta',\theta''}$ we obtain

$$g^{-1}_{n,\theta',\theta''}(y) = \frac{\ln y + n(c(\theta'') - c(\theta'))}{d(\theta'') - d(\theta')}, \quad y > 0. \quad (2.199)$$

Thus, for $n\in\Gamma^+$ the stopping bounds c_{n,θ_o,θ_1} and d_{n,θ_o,θ_1} are determined by

$$c_{n,\theta_o,\theta_1} = \frac{\ln B_n}{d(\theta_1) - d(\theta_o)} + \frac{n(c(\theta_1) - c(\theta_o))}{d(\theta_1) - d(\theta_o)} \quad (2.200)$$

and

$$d_{n,\theta_o,\theta_1} = \frac{\ln A_n}{d(\theta_1) - d(\theta_o)} + \frac{n(c(\theta_1) - c(\theta_o))}{d(\theta_1) - d(\theta_o)}, \quad (2.201)$$

respectively.

L e m m a 2.5.2. Let $\mathcal{P} = \{P_\theta,\ \theta \in (\underline{\theta}, \overline{\theta})\}$ be an MLRF. Let $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B_n, A_n\}_{n \in \Gamma^+}$ be an LRT for $H_0\colon\ \theta = \theta_0$ against $H_1\colon\ \theta = \theta_1$, $\underline{\theta} < \theta_0 < \theta_1 < \overline{\theta}$, and let $(N', \delta') = \{L_{n,\theta',\theta''}, B_n', A_n'\}_{n \in \Gamma^+}$ be an LRT for $H_0\colon\ \theta = \theta'$ against $H_1\colon\ \theta = \theta''$, $\underline{\theta} < \theta' < \theta'' < \overline{\theta}$. Then it is possible to choose sequences $\{B_n'\}_{n \in \Gamma^+}$ and $\{A_n'\}_{n \in \Gamma^+}$ in such a manner that

$$N = N' \quad \text{and} \quad \delta = \delta'. \tag{2.202}$$

This is implemented by

$$B_n' = g_{n,\theta',\theta''}(g^{-1}_{n,\theta_0,\theta_1}(B_n)),\ n \in \Gamma^+, \tag{2.203}$$

and

$$A_n' = g_{n,\theta',\theta''}(g^{-1}_{n,\theta_0,\theta_1}(A_n)),\ n \in \Gamma^+. \tag{2.204}$$

P r o o f. Suppose g_{n,θ_0,θ_1} is monotonically increasing. According to Lemma 2.5.1 the critical inequalities of test (N, δ) can be written as

$$c_{n,\theta_0,\theta_1} = g^{-1}_{n,\theta_0,\theta_1}(B_n) < T_n < g^{-1}_{n,\theta_0,\theta_1}(A_n) = d_{n,\theta_0,\theta_1},$$

$n \in \Gamma^+$. In a corresponding manner the critical inequalities of test (N', δ') can be written as

$$c_{n,\theta',\theta''} = g^{-1}_{n,\theta',\theta''}(B_n') < T_n < g^{-1}_{n,\theta',\theta''}(A_n') = d_{n,\theta',\theta''},$$

$n \in \Gamma^+$. Hence, we obtain (2.202) iff

$$c_{n,\theta',\theta''} = c_{n,\theta_0,\theta_1} \quad \text{and} \quad d_{n,\theta',\theta''} = d_{n,\theta_0,\theta_1}, \tag{2.205}$$

$n \in \Gamma^+$, which is equivalent to (2.203) and (2.204). Analogously, this lemma is established if g_{n,θ_0,θ_1} is monotonically decreasing. ∎

We discuss some consequences of this lemma, considering the subsequent example.

E x a m p l e 2.5.2. Continuation of Example 2.5.1. Let h be defined by

$$h = (d(\theta'') - d(\theta'))/(d(\theta_1) - d(\theta_0)). \tag{2.206}$$

Since d is strictly monotonically increasing in θ on $(\underline{\theta}, \overline{\theta})$, we obtain $h > 0$. Then by (2.198), (2.199), (2.203), (2.204) and (2.206) we obtain

$$B_n' = B_n^h \left(\exp(h(c(\theta_1) - c(\theta_0)) - (c(\theta'') - c(\theta'))\right)^n \tag{2.207}$$

and

$$A_n' = A_n^h \left(\exp(h(c(\theta_1) - c(\theta_0)) - (c(\theta'') - c(\theta'))\right)^n \tag{2.208}$$

for stopping bounds B_n' and A_n', $n \in \Gamma^+$, of test $(N', \delta') = \{L_{n,\theta',\theta''}, B_n', A_n'\}_{n \in \Gamma^+}$ considered in Lemma 2.5.2.

Some conclusions:

(i) Let $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$, $0 < B < 1 < A < \infty$. Then it follows from (2.207) and (2.208) for a given parameter pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$ with $\theta' < \theta''$ that, as a rule, the stopping bounds B_n' and A_n' for $n \in \Gamma^+$ of the equivalent test $(N', \delta') = \{L_{n,\theta',\theta''}, B_n', A_n'\}_{n\in\Gamma^+}$, proposed by Lemma 2.5.2, depend on parameters $\theta', \theta'', \theta_o$ and θ_1 as well as on the corresponding sampling stage n. Conversely, stopping bounds B_n' and A_n' do not depend on n if

$$h(c(\theta_1) - c(\theta_o)) - (c(\theta'') - c(\theta')) = 0. \tag{2.209}$$

Now, equations (2.206) and (2.209) are identical with equations (1.31) and (1.32), respectively. Hence, it follows from Lemma 1.6.3 for the considered exponential family that to any given WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ under the above assumptions the equivalent test $(N', \delta') = \{L_{n,\theta',\theta''}, B_n', A_n'\}_{n\in\Gamma^+}$ according to Lemma 2.5.2 is again a WLRT iff $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$ with $h > 0$. Then we have

$$B_n' = B^h \quad \text{and} \quad A_n' = A^h \quad \text{for} \quad n \in \Gamma^+.$$

(ii) Let $(N, \delta_c) = \{L_{n,\theta_o,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ be an LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ whose terminal decision rule δ_c is defined by

$$\delta_c = \chi\{L_{N,\theta_o,\theta_1} \geqslant c, N < \infty\}$$

for any given c, $0 < c < \infty$. By Theorem 2.2.1 this test is an MP-test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at level $E_{\theta_o}\delta_c$. According to Lemma 2.5.2 we can choose sequences $\{B_n'\}_{n\in\Gamma^+}$ and $\{A_n'\}_{n\in\Gamma^+}$ for every pair $\theta', \theta'' \in (\underline{\theta}, \overline{\theta})$, $\theta' < \theta''$, so that

$$N = N' = \begin{cases} \inf\{n \geqslant 1: L_{n,\theta',\theta''} \notin (B_n', A_n')\}, & \text{if such an n exists,} \\ \infty & \text{, otherwise.} \end{cases}$$

Analogously, we can obtain a sequence $\{k_n'\}_{n\in\Gamma^+}$ defined by

$$\begin{aligned} k_n' &= g_{n,\theta',\theta''}(g^{-1}_{n,\theta_o,\theta_1}(c)) \\ &= c^h \left(\exp(h(c(\theta_1) - c(\theta_o)) - (c(\theta'') - c(\theta')))\right)^n, \end{aligned} \tag{2.210}$$

$n \in \Gamma^+$. Hence, for every $n \in \Gamma^+$ inequality

$$L_{n,\theta_o,\theta_1} \geqslant c$$

is equivalent to

$$L_{n,\theta',\theta''} \geqslant k_n'.$$

If now δ_c' is a terminal decision rule defined by

$$\delta_c' = \chi\{L_{N',\theta',\theta''} \geq k_{N'}',\ N' < \infty\},$$

then we obtain $\delta_c = \delta_c'$. Indeed, as a rule, the critical number k_n' at stage $n \in \Gamma^+$ depends on n so that terminal decision rule δ_c' may have a structure which does not coincide with that of a terminal decision rule of an MP-test according to Theorem 2.2.1. Hence, test (N',δ_c') does not need to be also an MP-test for H_o: $\theta = \theta'$ against H_1: $\theta = \theta''$.

This seems to be one of the reasons which do not allow to obtain uniformly most powerful sequential LRTs. We notice that we obtain a sequence $\{k_n'\}_{n\in\Gamma^+}$ of critical numbers, defined by (2.210), which do not depend on $n \in \Gamma^+$ if $(\theta',\theta'') \overset{h}{\sim} (\theta_o,\theta_1)$ with $h > 0$ holds. Then this additional assumption ensures that test (N',δ_c') is also an MP-test for H_o: $\theta = \theta'$ against H_1: $\theta = \theta''$ at level $E_{\theta'}\delta_c'$ which corresponds with the statement of Corollary 2.2.2. ■

T h e o r e m 2.5.1. Let $\mathcal{P} = \{P_\theta,\ \theta \in (\underline{\theta},\bar{\theta})\}$ be an MLRF. Then every LRT $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B_n, A_n\}_{n\in\Gamma^+}$ for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ has a non-decreasing power function on $(\underline{\theta},\bar{\theta})$.

P r o o f. By Lemma 2.5.1 for every $\theta',\theta'' \in (\underline{\theta},\bar{\theta})$, $\theta' < \theta''$, sequences $\{B_n'\}_{n\in\Gamma^+}$ and $\{A_n'\}_{n\in\Gamma^+}$, $B_n' \leq A_n'$ for $n \in \Gamma^+$, exist so that test $(N',\delta') = \{L_{n,\theta',\theta''}, B_n', A_n'\}_{n\in\Gamma^+}$ satisfies

$$N = N' \quad \text{and} \quad \delta = \delta'. \tag{2.211}$$

By Theorem 2.3.1 test (N',δ') is an unbiased LRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta''$. This, together with (2.211), completes the proof. ■

We note that the monotonicity properties of the likelihood ratios L_{n,θ_o,θ_1} for $n \in \Gamma^+$ required in this theorem are sufficient but not necessary for the monotonicity of the power function. Alternative sufficient conditions for the monotonicity of the power function have been investigated by HOEL [43].

2.6 The termination property

In view of an implementation of a test (N,δ) only such tests are of importance which terminate with probability one for the parameters in consideration. Further we remark that the results obtained in the previous sections mainly concern these closed tests.

Here we will consider a criterion for the closedness of WLRT $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$. Since $\{N = \infty\} = \bigcap_{n=1}^{\infty}\{N > n\}$ and $\{\{N > n\}\}_{n\in\Gamma^+}$ is monotonically descending we have

$$P_\theta(N < \infty) = 1 \tag{2.212}$$

iff

$$P_\theta(N = \infty) = \lim_{n\to\infty} P_\theta(N > n) = 0.$$

For a WLRT we obtain

$$\begin{aligned} P_\theta(N > n) &= P_\theta(L_{1,\theta_o,\theta_1} \in (B,A), \ldots, L_{n,\theta_o,\theta_1} \in (B,A)) \\ &\leq P_\theta(L_{n,\theta_o,\theta_1} \in (B,A)) \\ &= 1 - P_\theta(L_{n,\theta_o,\theta_1} \leq B) - P_\theta(L_{n,\theta_o,\theta_1} \geq A),\ n \in \Gamma^+. \end{aligned}$$

Hence, a WLRT is closed w.r.t. θ if

$$\lim_{n\to\infty} P_\theta(L_{n,\theta_o,\theta_1} \in (B,A)) = 0$$

or

$$\lim_{n\to\infty} P_\theta(L_{n,\theta_o,\theta_1} \leq B) = 1 \text{ or } \lim_{n\to\infty} P_\theta(L_{n,\theta_o,\theta_1} \geq A) = 1.$$

respectively.

<u>L e m m a 2.6.1</u>. Let $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a WLRT for $H_o: \theta = \theta_o$ against $H_1: \theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density $f_\theta(x)$. If $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$ then

$$P_{\theta'}(Z_{1,\theta_o,\theta_1} = 0) < 1 \tag{2.213}$$

implies

$$P_{\theta'}(N < \infty) = 1. \tag{2.214}$$

P r o o f. Since $L_{n,\theta_o,\theta_1} \geq 0$ and $h \neq 0$ we obtain $L^h_{n,\theta_o,\theta_1} = 1$ iff $L_{n,\theta_o,\theta_1} = 1$. Hence, by $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$ and (2.213) we obtain

$$\begin{aligned} P_{\theta'}(L_{1,\theta',\theta''} = 1) &= P_{\theta'}(L^h_{1,\theta_o,\theta_1} = 1) = P_{\theta'}(L_{1,\theta_o,\theta_1} = 1) \\ &= P_{\theta'}(Z_{1,\theta_o,\theta_1} = 0) < 1. \end{aligned}$$

Applying Lemma 1.5.2 for $\theta_o = \theta'$ and $\theta_1 = \theta''$ we obtain

$$\lim_{n\to\infty} E_{\theta'} L^{\frac{1}{2}}_{n,\theta',\theta''} = 0$$

and by Lemma 1.5.1 and $(\theta', \theta'') \overset{h}{\sim} (\theta_o, \theta_1)$

$$P_{\theta'}(\lim_{n\to\infty} L_{n,\theta',\theta''} = 0) = P_{\theta'}(\lim_{n\to\infty} L^h_{n,\theta_o,\theta_1} = 0)$$

$$= P_{\theta'}(\lim_{n\to\infty} L_{n,\theta_o,\theta_1} = 0) = 1.$$

This implies

$$\lim_{n\to\infty} P_{\theta'}(L_{n,\theta_o,\theta_1} \leq B) = 1 \quad \text{for} \quad B > 0$$

and this is sufficient for (2.214) as stated above.■

If we are interested in an upper bound for the probability $P_\theta(N > n)$ of a WLRT we may use the following result by STEIN [74].

<u>L e m m a 2.6.2</u>. Let $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ be a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence of i.i.d. random variables having density $f_\theta(x)$. If

$$P_\theta(Z_{1,\theta_o,\theta_1} = 0) < 1 \tag{2.215}$$

then constants $0 < c < \infty$ and $0 < \varrho < 1$ exist so that

$$P_\theta(N > n) < c\varrho^n \tag{2.216}$$

for sufficiently large values of $n \in \Gamma^+$.

P r o o f. If $D^2_\theta Z_{1,\theta_o,\theta_1} = 0$ and (2.215) holds then (2.216) is evident. For $D^2_\theta Z_{1,\theta_o,\theta_1} \neq 0$ we consider the probability $P_\theta(N > mr)$ where m and r denote given integers. Since the $\{X_n\}_{n\in\Gamma^+}$ are i.i.d. also the increments $|Z_{k,\theta_o,\theta_1} - Z_{k-r,\theta_o,\theta_1}|$ are i.i.d. for k = r,2r,..., mr where $Z_{0,\theta_o,\theta_1} = 0$. Hence, for b = ln B and a = ln A we obtain

$$\begin{aligned} P_\theta(N > mr) &= P_\theta(Z_{k,\theta_o,\theta_1} \in (b,a) \text{ for } k = 1,\ldots,mr) \\ &\leq P_\theta(Z_{k,\theta_o,\theta_1} \in (b,a) \text{ for } k = r,2r,\ldots,mr) \\ &\leq P_\theta(|Z_{k,\theta_o,\theta_1} - Z_{k-r,\theta_o,\theta_1}| < a-b \text{ for } k = r,2r,\ldots,mr) \\ &= (P_\theta(|Z_{r,\theta_o,\theta_1}| < a-b))^m. \end{aligned} \tag{2.217}$$

Let Y_i be defined by $Y_i = \ln(f_{\theta_1}(X_i)/f_{\theta_o}(X_i))$, $i \in \Gamma^+$. Then the $\{Y_i\}_{i\in\Gamma^+}$ are also i.i.d. and we obtain

$$\begin{aligned} P_\theta(|Z_{r,\theta_o,\theta_1}| \geq a-b) &\geq P_\theta(\sum_{i=1}^{r} Y_i \geq a-b) \\ &\geq P_\theta(Y_i \geq \frac{a-b}{r} \text{ for } i = 1,\ldots,r) \\ &= (P_\theta(Y_1 \geq \frac{a-b}{r}))^r. \end{aligned} \tag{2.218}$$

In an analogous manner we obtain

$$P_\theta(|Z_{r,\theta_o,\theta_1}| \geq a-b) \geq (P_\theta(Y_i \leq -\frac{a-b}{r}))^r .$$

Furthermore, it follows from (2.215) that $P_\theta(Y_1 = 0) < 1$. Hence, there exists a constant $q > 0$ so that

$$P_\theta(Y_1 \geq \frac{a-b}{r}) > q \quad \text{or} \quad P_\theta(Y_1 \leq -\frac{a-b}{r}) > q \tag{2.219}$$

for sufficiently large values of r. Now, if $n = mr + k$ with $k \in \Gamma_o^+$, by (2.217) to (2.219) we obtain

$$P_\theta(N > n) \leq P_\theta(N > mr) < (1 - q^r)^m$$

and

$$(1 - q^r)^m = (1 - q^r)^{(1/r)(rm + k - k)} = c\varrho^n$$

if

$$\varrho = (1 - q^r)^{(1/r)} \quad \text{and} \quad c = \varrho^{-k}.$$

This completes the proof. ■

Since $\{N = \infty\} = \bigcap_{n=1}^{\infty}\{N > n\}$ and $\{\{N > n\}\}_{n\in\Gamma^+}$ is monotonically descending we obtain by (2.216)

$$P_\theta(N = \infty) = \lim_{n\to\infty} P_\theta(N > n) = \lim_{n\to\infty} c\varrho^n = 0$$

so that (2.216) ensures that N terminates with probability one. The property (2.216) will be described by saying that N is <u>exponentially bounded</u> w.r.t. θ. This notation, introduced by BERK [12], takes into account the fact that under the conditions of Lemma 2.6.2 a real number $t_o > 0$ exists so that $E_\theta\exp(tN) < \infty$ for every $t < t_o$ (see Lemma 2.7.1). Further aspects concerning the termination property of LRTs and exponentially bounded stopping times are considered by [13],[65],[68],[81] and [82].

2.7 The average sample number function

The average sample number function (ASN-function) is beside the power function one of the most important characteristics describing the statistical properties of a sequential test. If the test under consideration is a test based on a sequence $\{T_n\}_{n\in\Gamma^+}$ of statistics we may obtain assertions on the ASN-function by means of so-called moment equations. Moment equations are representation formulas for the moments involving beside expectation values of the variable Z_{N,θ_o,θ_1} also expectation values of statistics T_n, $n\in\Gamma^+$. WALD [77] used the so-called fundamental identity to obtain moment equations for the average sample size. These moment equations may be obtained

also in a direct manner applying methods of the computation of expectation values of randomly stopped sums (see e.g. CHOW et al. [22]).

We start here with a classical result concerning the existence of the moments of the sample size of WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables. Moreover, we will consider a moment equation for the first moment of the sample size of a WLRT which can be used to obtain lower bounds for the average sample size of a WLRT.

<u>Theorem 2.7.1</u>. Let $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ be a WLRT for $H_o: \theta = \theta_o$ against $H_1: \theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density $f_\theta(x)$. If $P_\theta(Z_{1,\theta_o,\theta_1} = 0) < 1$ then

$$E_\theta N^k < \infty \quad \text{for every } k\in\Gamma^+. \tag{2.220}$$

To establish this theorem we need the following result concerning the moment-generating function of N.

<u>Lemma 2.7.1</u>. Under the conditions of Theorem 2.7.1 a finite real number $t_o > 0$ exists with

$$E_\theta \exp(tN) < \infty \quad \text{for every } t < t_o. \tag{2.221}$$

Proof. Applying Lemma 2.6.2 we obtain

$$\begin{aligned}
E_\theta \exp(tN) &= \sum_{n=1}^{\infty} \exp(tn)P_\theta(N = n) \\
&= \sum_{m=0}^{\infty}\sum_{j=1}^{r} \exp((mr+j)t)\, P_\theta(N = mr + j) \\
&= \sum_{m=0}^{\infty} \exp(mrt) \sum_{j=1}^{r} \exp(jt)\; P_\theta(N = mr + j) \\
&\le \sum_{m=0}^{\infty} \exp(mrt)\; \exp(rt)\; P_\theta(N > mr) \\
&< \exp(rt) \sum_{m=0}^{\infty} \exp(mrt)\; c\varrho^{mr} \\
&= c\cdot\exp(rt) \sum_{m=0}^{\infty} (\exp(rt)\varrho^{r})^m, \qquad r\in\Gamma^+.
\end{aligned}$$

Like in the proof of Lemma 2.6.2 we obtain $0 < \varrho < 1$ and $c < \infty$ for sufficiently large values of r and the geometrical series

$$\sum_{m=0}^{\infty} (\exp(rt)\varrho^r)^m$$

converges iff $\exp(rt)\varrho^r < 1$. This is equivalent to $t < -\ln\varrho = t_o$ and we obtain (2.221). ■

Proof of Theorem 2.7.1: We note that for every $k \in \Gamma^+$ and $0 < t < t_o$ a finite $n' \in \Gamma^+$ exists with

$$n^k < \exp(nt) \quad \text{for every } n > n'.$$

Hence, we obtain

$$\begin{aligned} E_\theta N^k &= \sum_{n=1}^{\infty} n^k P_\theta(N = n) \\ &= \sum_{n=1}^{n'-1} n^k P_\theta(N = n) + \sum_{n=n'}^{\infty} n^k P_\theta(N = n) \\ &< (n' - 1)^k + \sum_{n=n'}^{\infty} \exp(nt)\, P_\theta(N = n) \\ &\leqslant (n' - 1)^k + E_\theta \exp(Nt) \end{aligned}$$

which is finite because of $n' < \infty$ and (2.221). ■

We now tend to a moment equation which is known as WALD's equation.

<u>Theorem 2.7.2</u>. Let N be a sample size based on the sequence of statistics $\{T_n\}_{n\in\Gamma^+}$. We suppose that the $\{T_n\}_{n\in\Gamma^+}$ are independent with the same mean $E_\theta T_1$ and $E_\theta |T_1| < \infty$. If $E_\theta N < \infty$ then we have

$$E_\theta T_N = E_\theta N \, E_\theta T_1 . \tag{2.222}$$

P r o o f. We follow WOLFOWITZ [83] and JOHNSON [47]. By the definition of T_N we have

$$\begin{aligned} T_N &= \sum_{n=1}^{\infty} \sum_{i=1}^{n} Y_i \chi\{N=n\} = \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} Y_i \chi\{N=n\} \\ &= \sum_{i=1}^{\infty} Y_i \sum_{n=i}^{\infty} \chi\{N=n\} = \sum_{i=1}^{\infty} Y_i \chi\{N \geqslant i\} . \end{aligned}$$

This implies

$$E_\theta T_N = \sum_{i=1}^{\infty} E_\theta Y_i \chi\{N \geqslant i\} = \sum_{i=1}^{\infty} E_\theta \chi\{N \geqslant i\} E_\theta(Y_i | \chi\{N \geqslant i\})$$

provided the interchange of expectation and summation is justified. We have:

(i) $\{N \geqslant i\} = \{N > i-1\}$ does not depend on $Y_i, Y_{i-1}, \ldots$ This implies $E_\theta(Y_i | \chi_{\{N \geqslant i\}}) = E_\theta Y_i = E_\theta Y_1$,

(ii) $E_\theta \chi_{\{N \geqslant i\}} = P_\theta(N \geqslant i)$,

(iii) $$\sum_{i=1}^{\infty} E_\theta \chi_{\{N \geqslant i\}} = \sum_{i=1}^{\infty} P_\theta(N \geqslant i) = \sum_{i=1}^{\infty} \sum_{n=i}^{\infty} P_\theta(N = n)$$

$$= \sum_{n=1}^{\infty} \sum_{i=1}^{n} P_\theta(N = n) = \sum_{n=1}^{\infty} n P_\theta(N = n) = E_\theta N.$$

Then (i), (ii) and (iii) imply $E_\theta T_N = E_\theta N \, E_\theta Y_1$. The interchange of expectation and summation is allowed if the series is absolutely convergent. Consider

$$\sum_{i=1}^{\infty} E_\theta \left| Y_i \chi_{\{N \geqslant i\}} \right| \leqslant \sum_{i=1}^{\infty} E_\theta |Y_i| E_\theta \chi_{\{N \geqslant n\}} = E_\theta |Y_1| \cdot E_\theta N.$$

This is finite because of $E_\theta |Y_1| < \infty$ and $E_\theta N < \infty$. Thus, the proof is complete. ■

By means of moment equation (2.222) we may obtain a lower bound for the average sample size of a closed WLRT.

L e m m a 2.7.2. Let $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$ be a WLRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence $\{X_n\}_{n \in \Gamma^+}$ of i.i.d. random variables having density $f_\theta(x)$. If $P_{\theta'}(Z_{1,\theta_0,\theta_1} = 0) < 1$, $E_{\theta'} |Z_{1,\theta_0,\theta_1}| < \infty$ and $(\theta', \theta'') \overset{h}{\sim} (\theta_0, \theta_1)$ then

$$E_{\theta'} N \geqslant \frac{(1 - M(\theta')) \ln \dfrac{1 - M(\theta'')}{1 - M(\theta')} + M(\theta') \ln \dfrac{M(\theta'')}{M(\theta')}}{h \cdot E_{\theta'} Z_{1,\theta_0,\theta_1}} . \qquad (2.223)$$

P r o o f. By Lemma 2.6.1 we obtain $E_{\theta'} N < \infty$. Applying Theorem 2.7.2 we have

$$E_{\theta'} Z_{N,\theta_0,\theta_1} = E_{\theta'} Z_{1,\theta_0,\theta_1} \cdot E_{\theta'} N. \qquad (2.224)$$

Otherwise, by Lemma 2.6.1 we obtain $P_{\theta'}(N < \infty) = 1$. This implies

$$E_{\theta'} Z_{N,\theta_0,\theta_1} = (1 - M(\theta')) E_{\theta'}(Z_{N,\theta_0,\theta_1} \,|\, H_0 \text{ is accepted})$$
$$+ M(\theta') E_{\theta'}(Z_{N,\theta_0,\theta_1} \,|\, H_1 \text{ is accepted}). \qquad (2.225)$$

By $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$, Jensens' inequality and Theorem 2.1.1 we obtain

$$E_{\theta'}(Z_{N,\theta_0,\theta_1} \mid H_0 \text{ is accepted}) = \frac{1}{h} E_{\theta'}(hZ_{N,\theta_0,\theta_1} \mid H_0 \text{ is acc.})$$

$$= \frac{1}{h} E_{\theta'}(\ln L^h_{N,\theta_0,\theta_1} \mid H_0 \text{ is accepted})$$

$$\leqslant \frac{1}{h} \ln E_{\theta'}(L^h_{N,\theta_0,\theta_1} \mid H_0 \text{ is accepted})$$

$$= \frac{1}{h} \ln \frac{1 - M(\theta'')}{1 - M(\theta')} . \qquad (2.226)$$

Analogously, we obtain

$$E_{\theta'}(Z_{N,\theta_0,\theta_1} \mid H_1 \text{ is accepted}) \leqslant \frac{1}{h} \ln \frac{M(\theta'')}{M(\theta')} . \qquad (2.227)$$

Hence, (2.224) to (2.227) imply

$$E_{\theta'}Z_{1,\theta_0,\theta_1} \cdot E_{\theta'}N \leqslant \frac{1}{h}\left((1- M(\theta')) \ln \frac{1 - M(\theta'')}{1 - M(\theta')} + M(\theta') \ln \frac{M(\theta'')}{M(\theta')}\right) \qquad (2.228)$$

Now, by $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ and Jensens' inequality we obtain

$$E_{\theta'}Z_{1,\theta_0,\theta_1} = \frac{1}{h} E_{\theta'}hZ_{1,\theta_0,\theta_1} = \frac{1}{h} E_{\theta'}Z_{1,\theta',\theta''}$$

$$= \frac{1}{h} E_{\theta'}\ln L_{1,\theta',\theta''}$$

$$\leqslant \frac{1}{h} \ln E_{\theta'}L_{1,\theta',\theta''} = \frac{1}{h} \ln 1 = 0.$$

Hence, dividing (2.228) by $E_{\theta'}Z_{1,\theta_0,\theta_1}$ we obtain (2.223). ∎

The bound given by (2,223) may not be the greatest lower bound for $E_{\theta'}N$. We refer in this context to HOEFFDING [41] and [42] who has derived certain further lower bounds for the average sample size.

<u>Corollary 2.7.1</u>. We suppose that Lemma 2.7.2 holds where

$$M(\theta_0) = \alpha \quad \text{and} \quad M(\theta_1) = 1 - \beta .$$

Then

$$E_{\theta_0}N \geqslant \frac{(1 - \alpha) \ln \frac{\beta}{1 - \alpha} + \alpha \ln \frac{1 - \beta}{\alpha}}{E_{\theta_0}Z_{1,\theta_0,\theta_1}} \qquad (2.229)$$

and

$$E_{\theta_1}N \geqslant \frac{\beta \ln \frac{\beta}{1 - \alpha} + (1 - \beta) \ln \frac{1 - \beta}{\alpha}}{E_{\theta_1}Z_{1,\theta_0,\theta_1}} . \qquad (2.230)$$

P r o o f. This corollary is an immediate conclusion of Lemma 2.72 $(\theta_o,\theta_1) \overset{1}{\sim} (\theta_o,\theta_1)$ and $(\theta_1,\theta_o) \overset{-1}{\sim} (\theta_o,\theta_1)$, respectively. ■

By means of inequalities (2.229) and (2.230) we may assess the expense which at least is required if we are interested in a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ with probabilities α and β of an error of first and second kind, respectively.

<u>C o r r l l a r y 2.7.2.</u> We suppose that Lemma 2.7.2 holds where

$$P_{\theta'}(L_{N,\theta_o,\theta_1} = B \mid H_o \text{ is acc.}) = P_{\theta'}(L_{N,\theta_o,\theta_1} = A \mid H_1 \text{ is acc.}) = 1 \tag{2.231}$$

Then

$$E_{\theta'}N = \frac{(1 - M^*(h)) \ln B + M^*(h) \ln A}{E_{\theta'} Z_{1,\theta_o,\theta_1}} \tag{2.232}$$

where $M^*(h)$ is defined by (2.21).

P r o o f. If (2.231) holds we have $M(\theta') = M^*(h)$, $E_{\theta'}(Z_{N,\theta_o,\theta_1} \mid H_o$ is accepted$) = \ln B$ and $E_{\theta'}(Z_{N,\theta_o,\theta_1} \mid H_1$ is accepted$) = \ln A$. This, together with (2.224) and (2.225), provides (2.232). ■

If instead of (2.231) we may only suppose that the excess at termination is small then (2.232) holds only approximately and the right-hand term of (2.232) is the so-called <u>WALD approximation for the ASN-function</u>.

2.8 The optimum property

The WLRT possesses a quite simple structure. Nevertheless, it possesses a surprising optimality property for the first time proved by WALD, WOLFOWITZ [78].

<u>T h e o r e m 2.8.1</u>. Let $(N, \delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ be a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence of i.i.d. random variables. Let $(\tilde{N}, \tilde{\delta})$ be any other test for these hypotheses. If the true risks $\alpha(\theta_o)$ and $\beta(\theta_1)$ of (N, δ) and $\tilde{\alpha}(\theta_o)$ and $\tilde{\beta}(\theta_1)$ of $(\tilde{N}, \tilde{\delta})$ satisfy

$$\tilde{\alpha}(\theta_o) \leq \alpha(\theta_o) \quad \text{and} \quad \tilde{\beta}(\theta_1) \leq \beta(\theta_1) \tag{2.233}$$

then

$$E_{\theta_o} N \leq E_{\theta_o} \tilde{N} \quad \text{and} \quad E_{\theta_1} N \leq E_{\theta_1} \tilde{N}. \tag{2.234}$$

The space of this booklet does not allow to present a complete proof of this theorem. The main idea of the classical proof of this theorem consists in the verification of the fact that a WLRT is a certain Bayes test in the sense of decision theory. We refer in this context to WALD, WOLFOWITZ [78], WOLFOWITZ [84], LEHMANN [53] and GHOSH [35] where proofs of this theorem are presented under the additional assumption of closedness of the considered tests. BURKHOLDER, WIJSMAN [20] have shown that this assumption is not necessary. Other optimum proofs have been given by MATTHES [58], SCHMITZ [67], LORDEN [54] and IRLE, SCHMITZ [45] where also the optimum property of a WLRT for processes with a continuous time parameter is established. We further refer to SCHMITZ [68].

The optimum property of a WLRT characterized by the above theorem is, of course, a pointwise optimum property which can be expressed also as follows. Among all tests whose error probabilities do not exceed those of the given WLRT this test possesses the smallest average sample size given under both hypotheses. In view of composite hypotheses or if a separating-parameter is given the optimum property of Theorem 2.8.1 can be extended by means of the conjugacy concept as follows.

T h e o r e m 2.8.2. Let $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ be a WLRT for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, $\underline{\theta} < \theta_0 < \theta_1 < \bar{\theta}$, $0 < B < 1 < A < \infty$, based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables. Suppose there exist parameters θ_0^* and θ_1^*, $\underline{\theta} < \theta_0^* < \theta_1^* < \bar{\theta}$, so that

(i) for every $\theta' \in (\underline{\theta},\theta_0^*)$ there exists a $\theta'' \in (\theta_1^*,\bar{\theta})$ where $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h>0$,

(ii) for every $\theta' \in (\theta_1^*,\bar{\theta})$ there exists a $\theta'' \in (\underline{\theta},\theta_0^*)$ where $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ with $h<0$.

Let $(\tilde{N},\tilde{\delta})$ be any other test for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. Denote by $M(\theta)$ and $\tilde{M}(\theta)$ the power function of (N,δ) and $(\tilde{N},\tilde{\delta})$, respectively. Then

$$\tilde{M}(\theta) \leq M(\theta) \quad \text{for} \quad \theta < \theta_0^* \tag{2.235}$$

and

$$\tilde{M}(\theta) \geq M(\theta) \quad \text{for} \quad \theta > \theta_1^* \tag{2.236}$$

imply

$$E_\theta N \leq E_\theta \tilde{N} \quad \text{for } \theta \in (\underline{\theta},\theta_0^*) \cup (\theta_1^*,\bar{\theta}). \tag{2.237}$$

P r o o f. (i) We suppose $\theta' \in (\underline{\theta},\theta_0^*)$. Then, there exists a $\theta'' \in (\theta_1^*,\bar{\theta})$ with $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ and $h>0$. This implies $L_{n,\theta',\theta''} = L^h_{n,\theta_0,\theta_1}$ for $n\in\Gamma^+$ and we obtain

$$(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+} = \{L^h_{n,\theta_0,\theta_1},B^h,A^h\}_{n\in\Gamma^+}$$

$$= \{L_{n,\theta',\theta''},B^h,A^h\}_{n\in\Gamma^+}$$

so that test (N,δ) is also a WLRT for $H_0: \theta = \theta'$ against $H_1: \theta = \theta''$ with stopping bounds $0 < B^h < 1 < A^h < \infty$. Applying Theorem 2.8.1 we obtain

$$E_\theta N < E_\theta \tilde{N} \quad \text{for} \quad \theta \in \{\theta',\theta''\}. \tag{2.238}$$

(ii) We suppose $\theta' \in (\theta_1^*,\bar{\theta})$. Then there exists a $\theta'' \in (\underline{\theta},\theta_0^*)$ with $(\theta',\theta'') \overset{h}{\sim} (\theta_0,\theta_1)$ and $h < 0$. This implies $(\theta'',\theta') \overset{-h}{\sim} (\theta_0,\theta_1)$ with $-h > 0$ and $L_{n,\theta',\theta''} = L^{-h}_{n,\theta_0,\theta_1}$ for every $n \in \Gamma^+$ so that

$$(N,\delta) = \{L_{n,\theta'',\theta'},B^{-h},A^{-h}\}_{n\in\Gamma^+}$$

holds. Applying Theorem 2.8.1 we again obtain (2.238). This completes the proof. ■

The conditions of this theorem are fulfilled, for instance, if we have a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having density

$$f_\theta(x) = h(x)\cdot\exp(d(\theta)t(x) - c(\theta)),\ x \in \mathfrak{X},\ \theta \in (\underline{\theta},\bar{\theta})$$

where Lemma 1.6.4 holds. Then, to given $\underline{\theta} < \theta_0 < \theta_1 < \bar{\theta}$ a uniquely determined separating-parameter $\theta^* \in (\underline{\theta},\bar{\theta})$ exists. Furthermore, for every $\theta_0^* < \theta^*$ there exists a corresponding $\theta_1^* > \theta^*$ so that

$$\zeta(\theta_0^*,\theta^*) = \zeta(\theta_1^*,\theta^*)$$

holds according to (1.57). Hence, we obtain a one-to-one correspondence between the elements of $(\underline{\theta},\theta_0^*)$ and $(\theta_1^*,\bar{\theta})$ in the sense of the conditions (i) and (ii) of Theorem 2.8.2. Then every test $(\tilde{N},\tilde{\delta})$ whose power function intersects the power function of (N,δ) between θ_0^* and θ_1^* as it is required by (2.235) and (2.236) possesses the optimum property (2.237). A special case arises if we choose $\theta_0^* = \theta^* = \theta_1^*$. Then, if $E_\theta N$ is a continuous function in θ on $(\underline{\theta},\bar{\theta})$ we obtain

$$E_\theta N \leq E_\theta \tilde{N} \quad \text{for} \quad \theta \in (\underline{\theta},\bar{\theta}).$$

under the conditions of Lemma 1.6.4.

3. The computation of the characteristics

In comparison with the large number of papers dealing with approximation methods for the computation of the power function or the average sample number function of a WLRT, we refer in this context also to the corresponding improvements of the WALD approximations for these characteristics given by KEMP [50], PAGE [60], BARACLOUGH, PAGE [9], there exists a comparatively small number of investigations concerning direct methods for the computation of the power function and the average sample size (see GHOSH [35], GOVINDARAJULU [37], JACKSON [46], JOHNSON4[48]). As a rule, direct methods are characterized by special assumptions about the probability distribution of the random variables considered there (see e.g. ALBERT [2], AROIAN, ROBINSON [6]). In those cases where the logarithm of the likelihood ratio Z_{1,θ_o,θ_1} is a random variable that only takes on a finite number of values which are the integer multiple of a given constant WALD [76], [77] used for the direct computation of the operating characteristic function the moment-generating function of Z_{1,θ_o,θ_1} and the fundamental identity. To obtain assertions about the operating characteristic function by this method we first have to determine the roots of a polynomial equation and second to solve a system of simultaneous linear equations. For the same situation GIRSHICK [36] has proposed an alternative method which involves solving a system of linear equations. The number of linear equations depends on the values of the stopping bounds and the constant mentioned above and may be very large in practical situations.
BURMAN [21], PÓLYA [64] and WALKER [79] considered exact formulas for a Bernoulli population. Finally, under the conditions of [36] and [76] it is possible to describe the operating characteristic function and the average sample size by means of difference equations (see e.g. GHOSH [35], Chapter 3.7) but the solution of these difference equations, in turn, requires the computation of the roots of a polynomial equation. All these methods do not allow to compute the operating characteristic function and the average sample size for truncated WLRTs.
Bartky [10], JONES [49] and BERNSTEIN [14] used the method of vector summation for the computation of the probabilities of acceptance. If some assumptions about the continuation region were fulfilled, they obtained assertions about the operating characteristic function and the average sample size after the inversion of a matrix. AROIAN [5] used an approach based on Markov chains. This approach is without

additional assumptions, e.g. a samll number of stages of the sampling plan or special assumptions concerning the underlying distribution, for the Bernoulli distribution we refer to [27], very laborious from the numerical point of view. HALD, MØLLER [38] used a similar approach for the design of two-, three- and seven stage sampling plans for the Poisson and Bernoulli distribution. Direct methods for the computation of the moments of the r^{th} order, $r \geq 2$, of the sample size seems to be unknown so far [34], [57].

By [28] a method has been developed allowing direct computation of the operating characteristic function, the power function and the moments of an arbitrary order of a WLRT based on a sequence of i.i.d. integer-valued random variables. There, we only have to suppose that the slope of the acceptance or rejection line of the WLRT is a rational number. This is a quite weak restriction because in practice we are always forced to use rational numbers. The class of distributions considered in [28] includes, for instance, the Bernoulli, Poisson, geometrical and negative binomial distribution. The amount of numerical calculations is comparatively small in comparison with the methods mentioned above. Mainly elementary vector operations are needed. The computation of more general characteristics of a WLRT by means of the method presented in [28] is investigated in [30]. An application example is described in [31] where the substitution of a sequential sampling plan for sampling by attributes by a sampling plan based on a sequence of Poisson distributed random variables is investigated.

Following [28] and [30] we will present a method for the computation of the characteristics of a WLRT based on a sequence of i.i.d. integer-valued random variables. In Section 3.1 we start with the investigation into some geometrical properties of the continuation region of the considered WLRT. In the subsequent section a direct method for the computation of the characteristics of a WLRT is developed under the assumption that the slope of the acceptance or rejection line of our WLRT is rational. In Sections 2.3 and 2.4 this method is applied to the computation of the power function and the moments of the sample size which can be reduced to that of solving of systems of linear equations which differ in their right-hand sides only. In Section 2.5 the computation of the distribution of the sample size of a WLRT is considered.

By means of the method for the computation of the power function in Section 2.2 it will be possible to obtain admissible WLRTs with a

sample size as small as possible. The corresponding procedure is described in Section 3.6.
The procedure of Section 3.2 is adapted to WLRTs based on a sequence of integer-valued random variables. Since one may discretize almost all statistical problems further application possibilities arise by suitable discretization of continuous problems. In connection with the so-called grouped observation a corresponding discretized sequential test procedure is considered in Section 3.7. We shall see that by means of the method presented there also test problems can be solved for which fixed sample counterparts are not known so that a further advantage of sequential tests over non-sequential tests will arise.
Moreover, an essential adventage of the method of Section 3.2 is that it can be extended to the computation of the characteristics of truncated WLRTs. For the power function and the moments of the sample size this will be done in Section 3.8.

Finally, we will consider in Section 3.9 a continuous inspection scheme for detecting a parameter change by means of a sequence of repeated WLRTs. Applying the methods of Sections 3.2 to 3.4 we will present a method for the computation of the average run length of such sampling schemes.

3.1 Properties of the continuation region

Consider a closed WLRT $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ for $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$, $0 < B < 1 < A < \infty$, based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of integer-valued i.i.d. random variables and suppose that

$$Z_{n,\theta_0,\theta_1} = \ln L_{n,\theta_0,\theta_1} = \gamma_1 \sum_{i=1}^{n} X_i - \gamma_0 n, \quad n\in\Gamma^+, \tag{3.1}$$

for any given $\gamma_0, \gamma_1 \in (-\infty,+\infty)$, $\gamma_1 \neq 0$. We generally assume that $\gamma_1 > 0$. In case of $\gamma_1 < 0$ the investigations are completely analogous.
Then critical inequalities

$$\ln B < Z_{n,\theta_0,\theta_1} < \ln A, \quad n\in\Gamma^+,$$

can be written as

$$\frac{\gamma_0}{\gamma_1} n + \frac{\ln B}{\gamma_1} < \sum_{i=1}^{n} X_i < \frac{\gamma_0}{\gamma_1} n + \frac{\ln A}{\gamma_1}, \quad n\in\Gamma^+. \tag{3.2}$$

We remark that our assumptions are fulfilled, for instance, if we have a sequence of i.i.d. random variables $\{X_n\}_{n\in\Gamma^+}$ with density

$$f_\theta(x) = h(x)\cdot \exp(d(\theta)\cdot x - c(\theta)),\ x \in \mathfrak{X} \subseteq \Gamma\ ,\ \theta \in \Theta \subseteq R^1, \quad (3.3)$$

where d is strictly monotonically increasing in θ on Θ. Then we obtain

$$\gamma_1 = d(\theta_1) - d(\theta_o) \quad \text{and} \quad \gamma_o = c(\theta_1) - c(\theta_o).$$

and the considered distribution class contains, e.g., the Bernoulli, Poisson, geometrical and negative binomial distribution. If θ^* denotes the separating-parameter given by (1.58) then we obtain

$$\frac{\gamma_o}{\gamma_1} = \frac{c(\theta_1) - c(\theta_o)}{d(\theta_1) - d(\theta_o)} = \frac{c'(\theta^*)}{d'(\theta^*)}$$

for family (3.3). Moreover, if $E_\theta X_1 = \theta_1$ we obtain even

$$\frac{\gamma_o}{\gamma_1} = \theta^*$$

since $E_\theta X_1 = c'(\theta)/d'(\theta)$.

We investigate some structural properties of the continuation region of test (N, δ) under consideration. Since the random variables $\{X_n\}_{n\in\Gamma^+}$ are assumed to be integer-valued the continuation region of our test can be described by means of a set of lattice points. Under the above assumptions let M be the set of lattice points given by

$$M = \left\{(n,k)\in\Gamma_o^+\times\Gamma : \frac{\gamma_o}{\gamma_1}n + \frac{\ln B}{\gamma_1} < k < \frac{\gamma_o}{\gamma_1}n + \frac{\ln A}{\gamma_1}\right\}, \quad (3.4)$$

then, for $n = 1,2,\ldots,$ we continue sampling as long as the points of the sequence $\left\{\left(n, \sum_{i=1}^{n} X_i\right)\right\}_{n\in\Gamma^+}$ belong to set M. If on stage n critical inequality (3.2) is violated for the first time we stop sampling and accept hypothesis H_o or H_1 if

$$\sum_{i=1}^{n} X_i \leqslant \frac{\gamma_o}{\gamma_1}n + \frac{\ln B}{\gamma_1} \quad \text{or} \quad \frac{\gamma_o}{\gamma_1}n + \frac{\ln A}{\gamma_1} \leqslant \sum_{i=1}^{n} X_i, \quad (3.5)$$

respectively. Hence, we obtain

$$N = \inf\left\{n \geqslant 1 : \left(n, \sum_{i=1}^{n} X_i\right) \notin M\right\}$$

and

$$\delta = \chi\left\{\sum_{i=1}^{N} X_i \geqslant \frac{\gamma_o}{\gamma_1}N + \frac{\ln A}{\gamma_1},\ N < \infty\right\}.$$

In view of this representation we may regard lattice point (0,0)

as the starting point for the test (N, δ) or the sequence of lattice points $\left\{\left(n, \sum_{i=1}^{n} X_i\right)\right\}_{n\in\Gamma^+}$. More general we may use also each other point of M as a starting point for a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$.

D e f i n i t i o n 3.1.1. Under the above assumptions we shall say a WLRT $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ based on the sequence $\{X_n\}_{n\in\Gamma^+}$ starts at the lattice point $(m,k)\in M$ if we proceed as follows:

(i) For $n = 1,2,\ldots$ we continue sampling as long as

$$\frac{\gamma_o}{\gamma_1}(m+n) + \frac{\ln B}{\gamma_1} < k + \sum_{i=1}^{n} X_{m+i} < \frac{\gamma_o}{\gamma_1}(m+n) + \frac{\ln A}{\gamma_1}. \tag{3.6}$$

(ii) We stop sampling on stage $n\in\Gamma^+$ and accept H_o or reject H_1, if on this stage

$$k + \sum_{i=1}^{n} X_{m+i} \leqslant \frac{\gamma_o}{\gamma_1}(m+n) + \frac{\ln B}{\gamma_1} \quad \text{or}$$

$$\frac{\gamma_o}{\gamma_1}(m+n) + \frac{\ln A}{\gamma_1} \leqslant k + \sum_{i=1}^{n} X_{m+i} \tag{3.7}$$

holds, respectively, for the first time.

We denote such a test by $T(m,k)$.

Denote by $N(m,k)$ and $\delta(m,k)$ the sample size and the terminal decision rule of test $T(m,k)$, respectively, then we have

$$N(m,k) = \inf\left\{n \geqslant 1: \left(m+n, k + \sum_{i=1}^{n} X_{m+i}\right) \notin M\right\}$$

and

$$\delta(m,k) = \chi\left\{k + \sum_{i=1}^{N(m,k)} X_{m+i} \geqslant \frac{\gamma_o}{\gamma_1}(m + N(m,k)) + \frac{\ln A}{\gamma_1}, N(m,k) < \infty\right\}.$$

Evidently, tests $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ and $T(0,0)$ are identically.

We remark that it is also possible to interprete test $T(m,k)$ as a conditional WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ based on $\{X_n\}_{n\in\Gamma^+}$ under the condition that we have reached lattice point $(m,k)\in M$ after m observations.

Before we start to compute the characteristics like, for instance, the power function or the moments of the sample size of test (N, δ)

we will consider some geometrical properties of M.

D e f i n i t i o n 3.1.2. Two lattice points $(m,k) \in M$ and $(m',k') \in M$ are said to be equivalent (write: $(m',k') \simeq (m,k)$) iff

$$k - \frac{\gamma_0}{\gamma_1} n = k' - \frac{\gamma_0}{\gamma_1} m'.$$

By this definition equivalent lattice points of M are characterized by the same distance to the straight line of acceptance $y(n) = (\gamma_0/\gamma_1)n + (\ln B)/\gamma_1$ or to the straight line of rejection $y(n) = (\gamma_0/\gamma_1)n + (\ln A)/\gamma_1$, respectively, taken in the direction of the ordinate.

The following lemma describes the situation where different equivalent lattice points will exist.

L e m m a 3.1.1. To any given $(m,k) \in M$ at least one further lattice point $(m',k') \in M$ exists with $(m,k) \simeq (m',k')$ and $m \neq m'$ iff γ_0/γ_1 is a rational number.

P r o o f. (i) Suppose $(m,k) \simeq (m',k')$, $m \neq m'$. Then $(m,k) \simeq (m',k')$ implies

$$\frac{\gamma_0}{\gamma_1} = \frac{k' - k}{m' - m}.$$

Since $m' \neq m$ and k, k', m and m' are integers γ_0/γ_1 is rational.
(ii) γ_0/γ_1 is rational. Then integers g_0 and $g_1 \neq 0$ exist where $\gamma_0/\gamma_1 = g_0/g_1$. Let (m',k') be defined by

$$m' = m + g_1 \quad \text{and} \quad k' = k + g_0.$$

Then we obtain

$$k' - \frac{\gamma_0}{\gamma_1} m' = k - \frac{\gamma_0}{\gamma_1} m + g_0 - g_1\frac{\gamma_0}{\gamma_1} = k - \frac{\gamma_0}{\gamma_1} m$$

and $m \neq m'$. This implies $(m,k) \simeq (m',k')$ with $m \neq m'$. ■

The importance of the existence of equivalent lattice points in M will become clear in the following section. There we shall see that tests which are started in equivalent lattice points will have the same characteristics.

3.2 The computation of the characteristics

Let $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$ be the WLRT considered in the previous section. In the sequel we will investigate the computation of the characteristics of this test if γ_0/γ_1 is a rational number. This task is embedded into the more general problem of the computation of

the caracteristics of test $T(m,k)$ for $(m,k) \in M$.

Denote by $S(m,k)$ for $(m,k) \in M$ the randomly stopped sum

$$S(m,k) = k + \sum_{i=1}^{N(m,k)} X_{m+i} .$$

Let further $Z(m,k)$ be a random variable defined by

$$Z(m,k) = g_1(k + \sum_{i=1}^{N(m,k)} X_{m+i}) - g_o(m + N(m,k))$$

$$= g_1 \, S(m,k) - g_o(m + N(m,k))$$

for $(m,k) \in M$. If g_o and g_1 are integers then $Z(m,k)$ is an integer-valued random variable. If we may suppose that

$$\frac{\gamma_o}{\gamma_1} = \frac{g_o}{g_1}, \quad g_o, g_1 \in \Gamma^+, \quad g_1 \neq 0, \tag{3.8}$$

we obtain the following relation between the likelihood ratio L_{N,θ_o,θ_1} and $Z(0,0)$.

$$\ln L_{N,\theta_o,\theta_1} = Z_{N,\theta_o,\theta_1} = \gamma_1 \sum_{i=1}^{N(0,0)} X_i - \gamma_o N(0,0)$$

$$= \gamma_1 \left(\sum_{i=1}^{N(0,0)} X_i - \frac{g_o}{g_1} N(0,0) \right)$$

$$= \frac{\gamma_1}{g_1} (g_1 \sum_{i=1}^{N(0,0)} X_i - g_o N(0,0))$$

$$= \frac{\gamma_1}{g_1} Z(0,0) . \tag{3.9}$$

Hence, under the assumptions (3.1) and (3.8) any characteristic of type

$$E_\theta w(N, L_{N,\theta_o,\theta_1}), \quad \theta \in \Theta,$$

can be reduced to the expectation value

$$E_\theta w\left(N(0,0), \exp\left(\frac{\gamma_1}{g_1} Z(0,0)\right)\right), \quad \theta \in \Theta.$$

For this reason, we will consider the computation of characteristics of test $T(m,k)$ which can be represented as expectation values of type

$$E_\theta w(N(m,k), Z(m,k)), \quad \theta \in \Theta,$$

for any given function $w: \Gamma_o^+ \times \Gamma \to R^1$. If we are able to calculate such expectation values we may also compute the corresponding characteristics of test $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n \in \Gamma^+}$.

L e m m a 3.2.1. Suppose that

$$\frac{\gamma_0}{\gamma_1} = \frac{g_0}{g_1}, \quad g_0, g_1 \in \Gamma, \quad g_1 > 0. \tag{3.10}$$

If $(m,k) \simeq (m',k')$ and $E_\theta w(N(m,k),Z(m,k))$ exists for $\theta \in \Theta$ then

$$E_\theta w(N(m,k),Z(m,k)) = E_\theta w(N(m',k'),Z(m',k')), \quad \theta \in \Theta. \tag{3.11}$$

P r o o f. Consider the distributions of $(N(m,k),Z(m,k))$ and $(N(m',k'),Z(m',k'))$. Since $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ is assumed to be closed here also tests $T(m,k)$, $(m,k) \in M$, are closed. By (3.10) critical inequalities

$$\frac{\ln B}{\gamma_1} + \frac{\gamma_0}{\gamma_1}(m + n) < k + \sum_{i=1}^{n} X_{m+i} < \frac{\ln A}{\gamma_1} + \frac{\gamma_0}{\gamma_1}(m + n), \quad n \in \Gamma^+,$$

are equivalent to

$$\frac{\gamma_0}{\gamma_0} \ln B < g_1(k + \sum_{i=1}^{n} X_{m+i}) - g_0(m + n) < \frac{g_0}{\gamma_0} \ln A, \quad n \in \Gamma^+.$$

Hence, we obtain

$$P_\theta(N(m,k) < \infty) = \sum_{n\in\Gamma^+} P_\theta(N(m,k)=n, Z(m,k) \notin (\frac{g_0}{\gamma_0} \ln B, \frac{g_0}{\gamma_0} \ln A))$$
$$= 1, \quad \theta \in \Theta.$$

This implies

$$P_\theta(N(m,k)=n, Z(m,k)=z) = 0 \quad \text{for} \quad (n,z) \in \Gamma^+ \times \Gamma$$

with $z \in (\frac{g_0}{\gamma_0} \ln B, \frac{g_0}{\gamma_0} \ln A)$. Analogously, we obtain

$$P_{\theta'}(N(m',k')=n, Z(m',k')=z) = 0 \quad \text{for} \quad (n,z) \in \Gamma^+ \times \Gamma$$

and $z \in (\frac{g_0}{\gamma_0} \ln B, \frac{g_0}{\gamma_0} \ln A)$. For $(n,z) \in \Gamma^+ \times \Gamma$ with $z \notin (\frac{g_0}{\gamma_0} \ln B, \frac{g_0}{\gamma_0} \ln A)$ we obtain

$$P_\theta(N(m,k)=n, Z(m,k)=z) = P_\theta(\frac{g_0}{\gamma_0} \ln B < g_1 \sum_{i=1}^{j} X_{m+i} - g_0 j + g_1 k - g_0 m$$
$$< \frac{g_0}{\gamma_0} \ln A \text{ for } j=1,\dots,n-1 \text{ and } g_1 \sum_{i=1}^{n} X_{m+i} - g_0 n + g_1 k - g_0 m = z) \tag{3.12}$$

and

$$P_\theta(N(m',k')=n, Z(m',k')=z) = P_\theta(\frac{g_0}{\gamma_0} \ln B < g_1 \sum_{i=1}^{j} X_{m'+i} - g_0 j + g_1 k'$$
$$- g_0 m' < \frac{g_0}{\gamma_0} \ln A \text{ for } j=1,\dots,n-1 \text{ and } g_1 \sum_{i=1}^{n} X_{m'+i} - g_0 n + g_1 k' - g_0 m' = z). \tag{3.13}$$

By $(m,k) \simeq (m',k')$ and (3.10) we obtain $g_1 k - g_o m = g_1 k' - g_o m'$. This, together with (3.12), (3.13) and the i.i.d.-property of $\{X_n\}_{n \in \Gamma^+}$, provides

$$P_\theta(N(m,k) = n, Z(m,k)=z) = P_\theta(N(m',k')=n, Z(m',k')=z)$$

also for $(n,z) \in \Gamma^+ \times \Gamma$ with $z \notin (\frac{g_o}{\gamma_o} \ln B, \frac{g_o}{\gamma_o} \ln A)$ which completes the proof. ■

An immediate conclusion of this lemma is that tests which start at equivalent lattice points will have the same characteristics. By means of this property we can obtain recursion formulas for the computation of the characteristics of tests $T(m,k)$, $(m,k) \in M$. Furthermore, depending on the structure of function w under consideration it will be even possible to modify these recursion formulas to systems of linear equations in certain cases.

We introduce the following notations:

$$k^{(o)}(m) = \min\left\{k \in \Gamma : k > \frac{\gamma_o}{\gamma_1} m + \frac{\ln B}{\gamma_1}\right\}, \quad m \in \Gamma_o^+;$$

$$k^{(1)}(m) = \max\left\{k \in \Gamma : k < \frac{\gamma_o}{\gamma_1} m + \frac{\ln A}{\gamma_1}\right\}, \quad m \in \Gamma_o^+;$$

$$K(m) = \left\{k \in \Gamma : k^{(o)}(m) \le k \le k^{(1)}(m)\right\}, \quad m \in \Gamma_o^+;$$

$$\bar{K}(m) = \Gamma - K(m), \quad m \in \Gamma_o^+;$$

$$w_k^\theta(m) = E_\theta w(N(m,k), Z(m,k)), \quad (m,k) \in M, \quad \theta \in \Theta;$$

$$w^\theta(m) = \left\{w_k^\theta(m)\right\}_{k \in K(m)};$$

$$w^\theta_{(1),k}(m) = E_\theta w(g_1 + N(m,k), Z(m,k)), \quad (m,k) \in M, \quad \theta \in \Theta;$$

$$w^\theta_{(1)}(m) = \left\{w^\theta_{(1),k}\right\}_{k \in K(m)};$$

$$C_{kk'}(m,m') = \Big\{k^{(o)}(j) \le k + \sum_{i=m+1}^{j} X_i \le k^{(1)}(j) \text{ for } j=m+1,\dots,m'-1 \text{ and }$$
$$k + \sum_{i=m+1}^{m'} X_i = k'\Big\}$$ - the event of reaching the lattice point $(m',k') \in \Gamma^+ \times \Gamma$ by test $T(m,k)$, $(m,k) \in M$, $m < m'$;

$$c^\theta_{kk'}(m,m') = P_\theta(C_{kk'}(m,m')), \quad \theta \in \Theta;$$

$$c^\theta(m,m') = \left\{c^\theta_{kk'}(m,m')\right\}_{k \in K(m), k' \in K(m')};$$

$c^{\theta}_{k'}(m,m') = \left\{ c^{\theta}_{kk'}(m,m') \right\}_{k\in K(m),\ k'\in \bar{K}(m')}$;

E - unit matrix of the same type as $C^{\theta}(m,m+g_1)$.

Then we obtain the following recursion formula.

<u>L e m m a 3.2.2.</u> Let $(N,\delta) = \left\{ L_{n,\theta_0,\theta_1}, B, A \right\}_{n\in\Gamma^+}$ be a closed WLRT based on a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. integer-valued random variables, where

$$Z_{n,\theta_0,\theta_1} = \gamma_1 \sum_{i=1}^{n} X_i - \gamma_0\, n, \quad n\in\Gamma^+,$$

$-\infty<\gamma_0<+\infty$, $0<\gamma_1<+\infty$. Suppose that

$$\frac{\gamma_0}{\gamma_1} = \frac{g_0}{g_1}, \quad g_0, g_1 \in \Gamma, \quad g_1 > 0. \tag{3.14}$$

Then

$$\vec{w}^{\theta}(m) = \vec{v}^{\theta}(m) + C^{\theta}(m,m+g_1)\, \vec{w}^{\theta}_{(1)}(m) \tag{3.15}$$

with

$$\vec{v}^{\theta}(m) = \sum_{n=1}^{g_1} \sum_{k'\in\bar{K}(m+n)} w(n, g_1 k' - g_0(m+n))\, \vec{c}^{\theta}_{k'}(m,m+n). \tag{3.16}$$

P r o o f. Consider the system of events

$$\left\{ \left\{C_{kk'}(m,m+1)\right\}_{k'\in\bar{K}(m+1)}, \dots, \left\{C_{kk'}(m,m+g_1)\right\}_{k'\in\bar{K}(m+g_1)}, \right.$$
$$\left. \left\{C_{kk'}(m,m+g_1)\right\}_{k'\in K(m+g_1)} \right\} \tag{3.17}$$

for $(m,k)\in M$ and $k\in K(m)$. This system forms a complete system of pairwise mutually and exclusive events. Then, by the formula of total probability, it follows

$$w^{\theta}_{k}(m) = \sum_{n=1}^{g_1} \sum_{k'\in\bar{K}(m+n)} E_{\theta}(w(N(m,k), Z(m,k)) \mid C_{kk'}(m,m+n)) \cdot$$
$$\cdot P_{\theta}(C_{kk'}(m,m+n)) + s'_k \tag{3.18}$$

with

$$s'_k = \sum_{k'\in K(m+g_1)} E_{\theta}(w(N(m,k), Z(m,k)) \mid C_{kk'}(m,m+g_1)) P_{\theta}(C_{kk'}(m,m+g_1)) \tag{3.19}$$

for every $k\in K(m)$. According to the definition of the events $C_{kk'}(m,m+n)$ we have

$$E_{\theta}(w(N(m,k), Z(m,k)) \mid C_{kk'}(m,m+n)) = w(n, g_1 k' - g_0(m+n)) \tag{3.20}$$

for $n = 1,\dots,g_1$, $k \in K(m)$ and $k' \in \bar{K}(m+n)$. Consider the conditional expectation values in the sum s_k' and put $k' = h + g_o$. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables applying Lemma 3.2.1 we obtain

$$E_\theta(w(N(m,k),Z(m,k)) \mid C_{kh+g_o}(m,m+g_1))$$

$$= E_\theta w(g_1+N(m+g_1,h+g_o),g_1S(m+g_1,h+g_o)-g_o(m+g_1+N(m+g_1,h+g_o)))$$

$$= E_\theta w(g_1+N(m+g_1,h+g_o),Z(m+g_1,h+g_o))$$

$$= E_\theta w(g_1+N(m,h),Z(m,h))$$

$$= E_\theta w^\theta_{(1),h}(m)$$

for $h \in K(m)$. This, together with (3.18), (3.19) and (3.20), provides (3.15). ■

Based on this lemma we may compute vector $\vec{w}^\theta(m)$ if in addition to the corresponding transition probabilities vector $\vec{w}^\theta_{(1)}(m)$ is known so that the computation of the characteristics $w^\theta_k(m)$ for $k \in K(m)$ can be reduced to the computation of the characteristics $w^\theta_{(1),k}(m)$ for $k \in K(m)$. In certain cases, the computation of $\vec{w}^\theta_{(1)}(m)$ will lead again to the computation of $\vec{w}^\theta(m)$ so that recursion formula (3.15) becomes a system of linear equations.

<u>T h e o r e m 3.2.1</u>. Suppose that Lemma 3.2.2 holds. If

$$w(g_1+n,z) = c\cdot w(n,z), \qquad c \in R^1, \tag{3.21}$$

then

$$(E - c\cdot C^\theta(m,m+g_1))\,\vec{w}^\theta(m) = \vec{v}^\theta(m) \tag{3.22}$$

where $\vec{v}^\theta(m)$ is determined by (3.16).

P r o o f. By (3.21) we have

$$w^\theta_{(1),k}(m) = c\cdot E_\theta w(N(m,k),Z(m,k)), \quad k \in K(m),$$

so that (3.22) is a conclusion of Lemma 3.2.2. ■

Under the conditions of this theorem expectation values of a function w depending only on Z(m,k) can be obtained as a solution of a system of linear equations. Corresponding examples are considered in Sections 3.3 and 3.4. For those cases where Theorem 3.2.1 can not be applied Lemma 3.2.2 can be modified as follows. According to Lemma 3.2.2 vector $\vec{w}^\theta(m)$ can be computed if vector $\vec{w}^\theta_{(1)}(m)$ is known. If for the computation of $\vec{w}^\theta_{(1)}(m)$ again Lemma 3.2.2 is used and if

we proceed in this manner we will obtain a sequence of recursion formulas. In doing this we still introduce the following notations:

$$w^{\theta}_{(r),k}(m) = E_{\theta}w(rg_1+N(m,k),Z(m,k)),\ r\in\Gamma^+_o,\ (m,k)\in M,\ \theta\in\Theta;$$

$$\vec{w}^{\theta}_{(r)}(m) = \{w^{\theta}_{(r),k}(m)\}_{k\in K(m)};$$

$$\vec{u}^{\theta}_{(r)}(m+n) = \sum_{k'\in\bar{K}(m+n)} w(rg_1+n,g_ok'-g_1(m+n))\vec{c}^{\,\theta}_{k'}(m),\ r,m\in\Gamma^+_o,\ n\in\Gamma^+,\ \theta\in\Theta.$$

$$\vec{v}^{\theta}_{(r)}(m) = \sum_{n=1}^{g_1}\vec{u}^{\theta}_{(r)}(m+n).$$

<u>Lemma 3.2.3</u>. Suppose that Lemma 3.2.1 holds. Then we have for $r = 0,1,2,\ldots$

(i) $$\vec{w}^{\theta}_{(r)}(m) = \vec{v}^{\theta}_{(r)}(m) + C^{\theta}(m,m+g_1)\,\vec{w}^{\theta}_{(r+1)}(m), \tag{3.23}$$

(ii) $$\vec{w}^{\theta}(m) = \sum_{j=1}^{r}(C^{\theta}(m,m+g_1))^j\,\vec{v}^{\theta}_{(j)}(m) + (C^{\theta}(m,m+g_1))^{r+1}\,\vec{w}^{\theta}_{(r+1)}(m), \tag{3.24}$$

(iii) $$\vec{w}^{\theta}(m) = \sum_{j=0}^{\infty}(C^{\theta}(m,m+g_1))^j\,\vec{v}^{\theta}_{(j)}(m). \tag{3.25}$$

Proof. Applying Lemma 3.2.2 to $w(rg_1+n,z)$ we obtain (3.23). Relation (3.24) is a conclusion of (3.23). Finally, (3.25) holds since $\vec{w}^{\theta}(m)$ is an expectation value. ■

The computation of $\vec{w}^{\theta}(m)$ according to Lemma 3.2.2, Theorem 3.2.1 or Lemma 3.2.3 requires the knowledge of transition matrix $C^{\theta}(m,m+g_1)$ for $m\in\Gamma^+_o$ as well as of transition vector $\vec{c}^{\,\theta}_{k'}(m,m+n)$ for $k'\in\bar{K}(m+n)$, $m\in\Gamma^+_o$, $n\in\Gamma^+$. Since sequence $\{X_n\}_{n\in\Gamma^+}$ is assumed to be a sequence of i.i.d. random variables we may state the following.

<u>Lemma 3.2.4</u>. Suppose that Lemma 3.2.2 holds.
(i) For $n = 1,2,\ldots,g_1$ we have

$$C^{\theta}(m,m+n) = \prod_{j=0}^{n-1} C^{\theta}(m+j,m+j+1) \tag{3.26}$$

where the elements of $C^{\theta}(m+j,m+j+1)$ are determined by

$$c^{\theta}_{kk'}(m+j,m+j+1) = P_{\theta}(X_1 = k' - k) \tag{3.27}$$

for $k\in K(m+j)$ and $k'\in K(m+j+1)$.

(ii) For $n = 1,2,\ldots,g_1$ and $k' \in \bar{K}(m+n)$ we have

$$\vec{c}^{\,\theta}_{k'}(m,m+n) = C^{\theta}(m,m+n-1)\ \vec{c}^{\,\theta}_{k'}(m+n-1,m+n) \tag{3.28}$$

where $C^{\theta}(m,m+n-1)$ is given by (3.26), where $C^{\theta}(m,m) = E$ and the elements of $c^{\theta}_{k'}(m+n-1,m+n)$ are determined by

$$c^{\theta}_{kk'}(m+n-1,m+n) = P_{\theta}(X_1 = k' - k) \tag{3.29}$$

for $k \in K(m+n-1)$.

P r o o f. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be integer-valued random variables this lemma follows immediately from the definition of transition probabilities $c^{\theta}_{kk'}(m,m+n)$. ■

That means we may parallel compute vectors $\vec{c}^{\,\theta}_{k'}(m,m+n)$ for $n = 1,\ldots,g_1$ to the computation of $C^{\theta}(m,m+g_1)$ by means of matrix multiplications. Since the one-step transition probabilities $c^{\theta}_{kk'}(m+j,m+j+1)$ and $c^{\theta}_{kk'}(m+n-1,m+n)$ can be reduced directly to single probabilities of random variable X_1 it is quite easy to generate the needed vectors $c^{\theta}_{k'}(m+n-1,m+n)$ and matrices $C^{\theta}(m+j,m+j+1)$. This effects that the amount of numerical calculations to obtain this quantities is comparatively small. Please note that on the other hand the amount of numerical calculations depends essentially on dimension d of vector $\vec{w}^{\,\theta}(m)$. We obtain in case of (3.14)

$$\begin{aligned} d(m) &= \operatorname{card} K(m) \\ &= \operatorname{card}\left\{k\in\Gamma : \frac{g_0}{g_1}\frac{\ln B}{\gamma_0} + \frac{g_0}{g_1} m < k < \frac{g_0}{g_1}\frac{\ln A}{\gamma_0} + \frac{g_0}{g_1} m\right\}. \end{aligned} \tag{3.30}$$

An approximation for d may be obtained by means of WALD's approximations for B and A. To given probabilities α and β, $0 < \alpha,\beta < 1$, $\alpha + \beta < 1$, for an error of first an second kind, respectively, we obtain

$$d(m) \approx d^{*}(m) = \operatorname{card}\left\{k\in\Gamma : \frac{g_0}{g_1}\frac{\ln\frac{\beta}{1-\alpha}}{\gamma_0} + \frac{g_0}{g_1} m < k < \frac{g_0}{g_1}\frac{\ln\frac{1-\beta}{\alpha}}{\gamma_0} + \frac{g_0}{g_1} m\right\} \tag{3.31}$$

3.3 The power function

We consider the computation of the power function of test T(m,k) by means of the method developed in the previous section. Additionally we introduce the following notations:

$m^{\theta}_{k}(m)$ - the probability of acceptance of H_1 by test T(m,k), $(m,k) \in M$, $\theta \in \Theta$;

$$\vec{m}^{\,\theta}(m) = \{m^{\theta}_{k}(m)\}_{k\in K(m)}; \tag{3.32}$$

$r_k^\Theta(m,m+n)$ - the probability of acceptance of H_1 by test $T(m,k)$ on n^{th} sampling stage, $n\in\Gamma^+$, $k\in K(m)$, $\Theta\in\Theta$;

$$\vec{r}^\Theta(m,m+n) = \{r_k^\Theta(m,m+n)\}_{k\in K(m)}.$$

Then we obtain the following assertion.

T h e o r e m 3.3.1. Suppose that Lemma 3.2.2 holds. Then for every $m\in\Gamma_o^+$ we have

$$(E - C^\Theta(m,m+g_1))\,\vec{m}^\Theta(m) = \sum_{n=1}^{g_1} \vec{r}^\Theta(m,m+n) \tag{3.33}$$

with

$$\vec{r}^\Theta(m,m+n) = C^\Theta(m,m+n-1)\,\vec{r}^\Theta(m+n-1,m+n) \tag{3.34}$$

and

$$r_k^\Theta(m+n-1,m+n) = P_\Theta(X_1 > k^{(1)}(m+n)-k) \tag{3.35}$$

for $n = 1,\dots,g_1$ and $k\in K(m+n-1)$.

P r o o f. By the definition of $m_k^\Theta(m)$ for every $k\in K(m)$ we have

$$m_k^\Theta(m) = E_\Theta\chi\{S(m,k) > k^{(1)}(m+N(m,k))\}$$

$$= E_\Theta\chi\{Z(m,k) > g_1k^{(1)}(m+N(m,k))-g_o(m+N(m,k))\}.$$

Hence, function w of Lemma 3.2.2 or Theorem 3.2.1 is given by

$$w(n,z) = \chi\{z > g_1k^{(1)}(m+n) - g_o(m+n)\}\ ,\quad (n,z)\in\Gamma_o^+\times\Gamma.$$

By (3.14) we obtain

$$k^{(1)}(m+n+g_1) = k^{(1)}(m+n) + g_o,\qquad n\in\Gamma_o^+.$$

This implies

$$w(n+g_1,z) = \chi\{z > g_1(k^{(1)}(m+n)+g_o)-g_o(m+n+g_1)\}$$

$$= w(n,z)$$

so that Theorem 3.2.1 can be applied. This implies (3.22) with $c = 1$. We consider the right-hand side of (3.22). For $k\in K(m)$ we have

$$v_k^\Theta(m) = \sum_{n=1}^{g_1}\ \sum_{k'\in\bar{K}(m+n)} c_{kk'}^\Theta(m,m+n)\,\chi\{g_1k'-g_o(m+n)>g_1k^{(1)}(m+n)-g_o(m+n)\}$$

$$= \sum_{n=1}^{g_1}\ \sum_{k'\in\bar{K}(m+n)} c_{kk'}^\Theta(m,m+n)\,\chi\{k' > k^{(1)}(m+n)\}$$

$$= \sum_{n=1}^{g_1} \sum_{k''\in K(m+n-1)} c^{\theta}_{kk''}(m,m+n-1) \sum_{k'>k^{(1)}(m+n)} c^{\theta}_{k''k'}(m+n-1,m+n)$$

$$= \sum_{n=1}^{g_1} \sum_{k''\in K(m+n-1)} c^{\theta}_{kk''}(m,m+n-1)\ P_{\theta}(k''+X_1 > k^{(1)}(m+n))$$

$$= \sum_{n=1}^{g_1} \sum_{k''\in K(m+n-1)} c^{\theta}_{kk''}(m,m+n-1)\ r^{\theta}_{k''}(m+n-1,m+n)\ .$$

This implies

$$\vec{v}^{\theta}(m) = \sum_{n=1}^{g_1} C^{\theta}(m,m+n-1)\ \vec{r}^{\theta}(m+n-1,m+n)$$

for $n = 1,\dots,g_1$ and, together with Theorem 3.2.1, (3.34) and (3.35), this completes the proof.■

We remark that according to Lemma 3.2.1 relation (3.34) can be written also as

$$\vec{r}^{\theta}(m,m+n) = \prod_{j=0}^{n-2} C^{\theta}(m+j,m+j+1)\ \vec{r}^{\theta}(m+n-1,m+n) \tag{3.36}$$

for $n = 2,\dots,g_1$. That means that the right-hand side of (3.33) can be again computed parallel to the computation of $C^{\theta}(m,m+g_1)$. If we are interested in the power function $M(\theta)$ of $(N,\delta) = \{L_{n,\theta_0,\theta_1}, B,A\}_{n\in\Gamma^+}$ we have to compute the power function of $T(0,0)$. If Theorem 3.3.1 holds we have

$$M(\theta) = m^{\theta}_{0}(0), \quad \theta \in \Theta. \tag{3.37}$$

In an analogous manner we are able to compute the operating characteristic function of $T(m,k)$. Denote by

$q^{\theta}_{k}(m)$ - the probability of the acceptance of H_0 by test $T(m,k)$, $(m,k)\in M$, $\theta\in\Theta$;

$\vec{q}^{\theta}(m) = \{q^{\theta}_{k}(m)\}_{k\in K(m)}$; (3.38)

$a^{\theta}_{k}(m,m+n)$ - the probability of acceptance of H_0 by test $T(m,k)$ on the n^{th} sampling stage, $n\in\Gamma^+$, $k\in K(m)$, $\theta\in\Theta$;

$\vec{a}^{\theta}(m,m+n) = \{a^{\theta}_{k}(m,m+n)\}_{k\in K(m)}$.

Then the following theorem holds.

Theorem 3.3.2. Suppose that Lemma 3.2.2 holds. Then, for every $m \in \Gamma_o^+$, we have

$$(E - C^\theta(m,m+g_1))\, \vec{q}^{\,\theta}(m) = \sum_{n=1}^{g_1} \vec{a}^{\,\theta}(m,m+n) \tag{3.39}$$

with

$$\vec{a}^{\,\theta}(m,m+n) = C^\theta(m,m+n-1)\, \vec{a}^{\,\theta}(m+n-1,m+n) \tag{2.40}$$

and

$$a_k^\theta(m+n-1,m+n) = P_\theta(X_1 < k^{(o)}(m+n)-k) \tag{2.41}$$

for $n = 1,\dots,g_1$ and $k \in K(m+n-1)$.

The computation of the power function and the operating characteristic function according to Theorem 3.3.1 and 3.3.2, respectively, by solving systems of linear equations requires that ratio γ_o/γ_1 is rational. If this assumption is not fulfilled we may obtain two-sided bounds for the power function of test $(N,\delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ if we proceed as follows.

Let g_o^*, g_1^*, g_o^{**} and g_1^{**} be integers so that

$$\frac{g_o^*}{g_1^*} \le \frac{\gamma_o}{\gamma_1} \le \frac{g_o^{**}}{g_1^{**}}, \tag{3.42}$$

$g_1^* > 0$ and $g_1^{**} > 0$. Let M be the continuation region defined by

$$M^* = \left\{(n,k)\in\Gamma_o^+\times\Gamma : \frac{g_o^*}{g_1^*}n + \frac{\ln B}{\gamma_1} < k < \frac{g_o^*}{g_1^*}n + \frac{\ln A}{\gamma_1}\right\}$$

and, according to the Definition 3.2.1, let $T^*(m,k)$ be for this continuation region a WLRT for $H_o: \theta = \theta_o$ against $H_1: \theta = \theta_1$ with start point $(m,k)\in M^*$. Denote by $m_k^{*\theta}(m)$ the power function of $T^*(m,k)$. In an analogous manner to the given integers g_o^{**} and g_1^{**} we define a continuation region M^{**} and tests $T^{**}(m,k)$ with power function $m_k^{**\theta}(m)$ for $(m,k)\in M^{**}$. Then, since g_o^*, g_1^*, g_o^{**} and g_1^{**} are integers power functions $m_k^{*\theta}(m)$ for $(m,k)\in M^*$ and $m_k^{**\theta}(m)$ for $(m,k)\in M^{**}$ can be computed applying Theorem 3.3.1. By (3.42) we obtain

$$\frac{g_o^*}{g_1^*}n + \frac{\ln B}{\gamma_1} \le \frac{\gamma_o}{\gamma_1}n + \frac{\ln B}{\gamma_1} \le \frac{g_o^{**}}{g_1^{**}}n + \frac{\ln B}{\gamma_1}$$

and

$$\frac{g_o^*}{g_1^*}n + \frac{\ln A}{\gamma_1} \le \frac{\gamma_o}{\gamma_1}n + \frac{\ln A}{\gamma_1} \le \frac{g_o^{**}}{g_1^{**}}n + \frac{\ln A}{\gamma_1}$$

for every $n\in\Gamma^+$ and the acceptance of H_1 by $T^{**}(0,0)$ implies the acceptance of H_1 by (N,δ). This further implies the acceptance of H_1 by $T^*(0,0)$. Then, for the power function $M(\theta)$ of (N,δ) we obtain

$$m_o^{**\theta}(0) \leq M(\theta) \leq m_o^{*\theta}(0), \quad \theta \in \Theta.$$

An analogous assertion can be obtained for the operating characteristic function.

3.4 The moments of the sample size

We consider the computation of the moments of the sample size of test $T(m,k)$ by means of the method of Section 3.3. We still introduce the following notations.

$e^{\theta}_{r,k}(m) = E_{\theta}N^{r}(m,k)$ – the r^{th} moment of sample size $N(m,k)$ of test $T(m,k)$, $(m,k) \in M$, $\theta \in \Theta$, $r \in \Gamma_o^+$;

$\vec{e}^{\,\theta}_{r}(m) = \{e^{\theta}_{r,k}(m)\}_{k \in K(m)}$.

Theorem 3.4.1. Suppose that Lemma 3.2.2 holds where $D^2_{\theta}X_1 > 0$. Then for every $r \in \Gamma_o^+$ we have

$$(E - C^{\theta}(m,m+g_1))\,\vec{e}^{\,\theta}_{r}(m) = \sum_{n=1}^{g_1} n^{r}(\vec{a}^{\,\theta}(m,m+n) + \vec{r}^{\,\theta}(m,m+n))$$
$$+ \sum_{j=0}^{r-1} \binom{r}{j} g_1^{r-j} C^{\theta}(m,m+g_1)\,\vec{e}^{\,\theta}_{j}(m) \qquad (3.43)$$

where $\vec{a}^{\,\theta}(m,m+n)$ and $\vec{r}^{\,\theta}(m,m+n)$ are determined by (3.40) and (3.34), respectively.

Proof. Assumption $D^2_{\theta}X_1 > 0$ provides $E_{\theta}N^{r}(m,k) < \infty$ for every $r \in \Gamma_o^+$. Applying Lemma 3.2.3 we have

$$w(n,z) = n^{r}, \qquad (n,z) \in \Gamma_o^+ \times \Gamma,$$

and we obtain

$$v^{\theta}_{k}(m) = \sum_{n=1}^{g_1} \sum_{k' \in \bar{K}(m+n)} n^{r} \cdot c^{\theta}_{kk'}(m,m+n)$$
$$= \sum_{n=1}^{g_1} n^{r}(a^{\theta}_{k}(m,m+n) + r^{\theta}_{k}(m,m+n)) \qquad (3.44)$$

for $k \in K(m)$. Furthermore, we have

$$w^{\theta}_{(1),k}(m) = E_{\theta}(g_1 + N(m,k))^{r}$$
$$= \sum_{j=0}^{r} \binom{r}{j} g_1^{r-j}\, E_{\theta}N^{j}(m,k)$$

$$= \sum_{j=0}^{r-1} \binom{r}{j} g_1^{r-j} \, e^{\theta}_{j,k}(m) + e^{\theta}_{r,k}(m) \tag{3.45}$$

for $k \in K(m)$. Then, by Lemma 3.2.2, (3.16) and (3.45), we obtain (3.43). ■

To illustrate this theorem we consider the computation of the first and second moment of sample size of test $T(m,k)$.

<u>Corollary 3.4.1</u>. Suppose that Theorem 3.4.1 holds. Then we have

(i) $$(E - C^{\theta}(m,m+g_1)) \, \vec{e}^{\,\theta}_1(m) = \sum_{n=1}^{g_1} n(\vec{a}^{\,\theta}(m,m+n) + \vec{r}^{\,\theta}(m,m+n)) + g_1 C^{\theta}(m,m+g_1) \, \vec{1}, \tag{3.46}$$

(ii) $$(E - C^{\theta}(m,m+g_1)) \, \vec{e}^{\,\theta}_2(m) = \sum_{n=1}^{g_1} n^2(\vec{a}^{\,\theta}(m,m+n) + \vec{r}^{\,\theta}(m,m+n)) + g_1^2 C^{\theta}(m,m+g_1) \, \vec{1} + 2g_1 C^{\theta}(m,m+g_1) \, \vec{e}^{\,\theta}_1(m) \tag{3.47}$$

where $\vec{1} = \{1\}_{k \in K(m)}$.

That means that based on Theorem 3.4.1 we may compute successively, beginning with average sample sizes $e^{\theta}_{1,k}(m)$ for $k \in K(m)$, the moments of sample size $e^{\theta}_{2,k}(m)$, $e^{\theta}_{3,k}(m), \ldots$ for $k \in K(m)$. In doing this, we have step by step to solve systems of linear equations which differ only in their right-hand sides. If we are again interested in the moments of sample size $E_{\theta}N^r$, $r \in \Gamma^+_o$, of test $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n \in \Gamma^+}$ we have to compute the corresponding moments of the sample size of test $T(0,0)$. If Theorem 3.4.1 holds we have

$$E_{\theta}N^r = e^{\theta}_{r,0}(0), \quad \theta \in \Theta.$$

<u>3.5 The distribution of the sample size</u>

In view of a truncation of a WLRT the distribution of the sample size will play a significant role. Here we will consider the computation of the distribution of sample size of test $T(m,k)$. Denote by

$$p^{\theta}_k(m;n) = P_{\theta}(N(m,k) = n), \quad n \in \Gamma^+, \tag{3.48}$$

the probability that test $T(m,k)$ terminates on sampling stage n, $(m,k) \in M$, $\theta \in \Theta$.

Lemma 3.5.1. Suppose that Lemma 3.2.2 holds.

(i) For $n = 1,2,\ldots,g_1$ we have

$$\vec{p}^{\theta}(m;n) = C^{\theta}(m,m+n-1)(\vec{a}^{\theta}(m+n-1,m+n) + \vec{r}^{\theta}(m+n-1,m+n)). \quad (3.49)$$

(ii) For $n = rg_1 + s$ with $r\in\Gamma^+$ and $s = 1,\ldots,g_1$ we have

$$\vec{p}^{\theta}(m;n) = (C^{\theta}(m,m+g_1))^r\, \vec{p}^{\theta}(m;s). \quad (3.50)$$

P r o o f. By the definition of $p_k^{\theta}(m;n)$ we have

$$p_k^{\theta}(m;n) = E_{\theta}(N(m,k) = n) = E_{\theta}\chi_{\{N(m,k) = n\}},\quad k\in K(m),$$

so that function w of Lemma 3.2.2 is given by

$$w(n',z) = \chi_{\{n' = n\}},\qquad (n',z)\in\Gamma_o^+\times\Gamma .$$

Then by Lemma 3.2.2 we obtain

$$p_k^{\theta}(m;n) = \sum_{n'=1}^{g_1}\ \sum_{k'\in\bar{K}(m+n')}\chi_{\{n' = n\}}\, c_{kk'}^{\theta}(m,m+n')$$

$$+ \sum_{k'\in K(m+g_1)} c_{kk'}^{\theta}(m,m+g_1)E_{\theta}\chi_{\{g_1+N(m,k) = n\}},\quad k\in K(m). \quad (3.51)$$

(i) Suppose $n = 1,\ldots,g_1$: Since $N(m,k)\geqslant 1$ we have $\chi_{\{g_1+N(m,k)=n\}} = 0$. Hence, (3.51) yields

$$p_k^{\theta}(m;n) = \sum_{k'\in\bar{K}(m+n)} c_{kk'}^{\theta}(m,m+n)$$

$$= \sum_{k''\in K(m+n-1)} c_{kk''}^{\theta}(m,m+n-1)(a_{k''}^{\theta}(m+n-1,m+n)+r_{k''}^{\theta}(m+n-1,m+n))$$

for $k\in K(m)$ which establishes (3.49).

(ii) Suppose $n = g_1 + s$, $s = 1,\ldots,g_1$: Then for $n = 1,\ldots,g_1$ we have $\chi_{\{n' = n\}} = 0$. Hence, instead of (3.51) we obtain

$$p_k^{\theta}(m;n) = p_k^{\theta}(m;g_1+s)$$

$$= \sum_{k'\in K(m+g_1)} c_{kk'}^{\theta}(m,m+g_1)E_{\theta}\chi_{\{g_1+N(m,k) = g_1+s\}}$$

$$= \sum_{k'\in K(m+g_1)} c_{kk'}^{\theta}(m,m+g_1)\, p_k^{\theta}(m;s)$$

for $k \in K(m)$ so that (3.50) holds for $r = 1$. Repeating this step (3.50) can be established for $r = 2,3,\ldots$ ■

3.6 Admissible tests

A test (N, δ) for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ is said to be admissible at size (α, β) if to given probabilities α and β its power function $M(\theta)$ satisfies

$$M(\theta_0) \leqslant \alpha \quad \text{and} \quad M(\theta_1) \geqslant 1 - \beta .$$

In Section 2.4 we have investigated certain possibilities to obtain values for stopping bounds B and A of WLRT $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$. In general we will obtain an admissible WLRT if

$$B = \beta \quad \text{and} \quad A = 1/\alpha .$$

Indeed, in this case B is less and A is greater than necessary as a rule. In view of a sample size as small as possible difference A - B should be chosen as small as possible. In the sequel we will present a procedure which follows this requirement for the class of tests considered in Section 3.2.

Let $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n \in \Gamma^+}$ be a WLRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on the sequence of i.i.d. integer-valued random variables $\{X_n\}_{n \in \Gamma^+}$. Like in Section 3.2 we suppose that

$$Z_{n,\theta_0,\theta_1} = \ln L_{n,\theta_0,\theta_1} = \gamma_1 \sum_{i=1}^{n} X_i - \gamma_0 n, \quad n \in \Gamma^+, \tag{3.52}$$

holds for any given real numbers γ_0 and $\gamma_1 > 0$. Further we suppose that

$$\frac{\gamma_0}{\gamma_1} = \frac{g_0}{g_1}, \quad g_0, g_1 \in \Gamma^+, \quad g_1 > 0, \tag{3.53}$$

and obtain

$$Z_{n,\theta_0,\theta_1} = \gamma_1 \Big(\sum_{i=1}^{n} X_i - \frac{g_0}{g_1} n \Big), \quad n \in \Gamma^+,$$

and

$$\frac{g_1}{\gamma_1} Z_{n,\theta_0,\theta_1} = g_1 \sum_{i=1}^{n} X_i - g_0 n, \quad n \in \Gamma^+,$$

so that random variables $(g_1/\gamma_1) Z_{n,\theta_0,\theta_1}$, $n \in \Gamma^+$, are integer-valued random variables.

For $z, s \in \Gamma_0^+$, $0 < z < s$, let $T_{z,s}$ be a further WLRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ with sample size

$$N_{z,s} = \inf\left\{n \geqslant 1: \frac{g_1}{\gamma_1} Z_{n,\theta_o,\theta_1} + z \notin (0,s)\right\} \tag{3.54}$$

and terminal decision rule

$$\delta_{z,s} = \chi_{\left\{\frac{g_1}{\gamma_1} Z_{N_{z,s},\theta_o,\theta_1} + z \geqslant s,\ N_{z,s} < \infty\right\}}. \tag{3.55}$$

Then we have

$$(N_{z,s}, \delta_{z,s}) = \left\{L_{n,\theta_o,\theta_1}, \exp(-\frac{\gamma_1}{g_1} z), \exp(\frac{\gamma_1}{g_1}(s-z))\right\}$$

and test $(N_{z,s}, \delta_{z,s})$ coincides with test (N, δ) if

$$B = \exp(-\frac{\gamma_1}{g_1} z) \quad \text{and} \quad A = \exp(\frac{\gamma_1}{g_1}(s-z)).$$

Conversely, for $z,s \in \Gamma^+$, $z < s$, this relation provides values for stopping bounds B and A of $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ so that (N, δ) coincides with $(N_{z,s}, \delta_{z,s})$.

Denote by $M_{z,s}(\theta)$, $\theta \in \Theta$, the power function of $(N_{z,s}, \delta_{z,s})$. Then for every fixed $\theta \in \Theta$ and $s \in \Gamma^+$ $M_{z,s}(\theta)$ is a non-decreasing function in z on $\{1,2,\ldots,s-1\}$. Between power functions $m_k^{\theta}(m)$ of test $T(m,k)$ for $(m,k) \in M$ and power functions $M_{z,s}(\theta)$ of tests $T_{z,s}$ defined above the following connection consists.

<u>L e m m a 3.6.1</u>. Suppose that $(N, \delta) = \{L_{n,\theta_o,\theta_1}, B, A\}_{n\in\Gamma^+}$ satisfies (3.52) and (3.53). Let M be the set of lattice points defined by (3.4). Let s be an integer given by

$$s = -\,\text{entier}(\frac{g_1}{\gamma_1} \ln B) - \text{entier}(-\frac{g_1}{\gamma_1} \ln A). \tag{3.56}$$

Then, for every $(m,k) \in M$ we have

$$M_{z,s}(\theta) = m_k^{\theta}(m), \quad \theta \in \Theta, \tag{3.57}$$

with

$$z = g_1 k - g_o m - \text{entier}(\frac{g_1}{\gamma_1} \ln B). \tag{3.58}$$

Especially, if $(m,k) \simeq (m',k')$ we have again

$$M_{z,s}(\theta) = m_{k'}^{\theta}(m'), \quad \theta \in \Theta. \tag{3.59}$$

P r o o f. Test $T_{z,s}$ is characterized by critical inequalities

$$0 < \frac{g_1}{\gamma_1} Z_{n,\theta_o,\theta_1} + z < s, \quad n \in \Gamma^+,$$

or

$$0 < g_1 \sum_{i=1}^{n} X_i - g_o n + z < s, \quad n \in \Gamma^+, \tag{3.60}$$

respectively. Test $T(m,k)$ is characterized by critical inequalities

$$\frac{g_o}{g_1}(m+n) + \frac{g_o}{g_1}\frac{\ln B}{\gamma_o} < k + \sum_{i=1}^{n} X_{m+i} < \frac{g_o}{g_1}(m+n) + \frac{g_o}{g_1}\frac{\ln A}{\gamma_o}, \quad n \in \Gamma^+.$$

Since the $\{X_n\}_{n\in\Gamma^+}$ are discrete random variables these inequalities are equivalent to

$$\operatorname{entier}\left(\frac{g_o}{\gamma_o}\ln B\right) < g_1 \sum_{i=1}^{n} X_{m+i} - g_o n + g_1 k - g_o m$$

$$< -\operatorname{entier}\left(-\frac{g_o}{\gamma_o}\ln A\right), \quad n \in \Gamma^+. \tag{3.61}$$

We note that (3.53) implies $g_o/\gamma_o = g_1/\gamma_1$. Hence, (3.61) implies

$$0 < g_1 \sum_{i=1}^{n} X_{m+i} - g_o n + z < s, \quad n \in \Gamma^+, \tag{3.62}$$

where z and s are defined by (3.58) and (3.56), respectively. Comparing (3.60) and (3.62) we obtain (3.57) by means of the i.i.d.-property of $\{X_n\}_{n\in\Gamma^+}$. The critical inequalities of test $T(m',k')$ can be written as

$$0 < g_1 \sum_{i=1}^{n} X_{m'+i} - g_o n + g_1 k' - g_o m' - \operatorname{entier}\left(\frac{g_o}{\gamma_o}\ln B\right) < s, \tag{3.63}$$

$n \in \Gamma^+$. By $(m,k) \simeq (m',k')$ and (3.53) we obtain

$$g_1 k - g_o m = g_1 k' - g_o m'.$$

This, together with (3.63), (3.58), (3.62) and the i.i.d.-property of $\{X_n\}_{n\in\Gamma^+}$, provides (3.59).■

Hence, the computation of power functions $M_{z,s}(\Theta)$ for $z \in \{1,\ldots,s-1\}$ can be reduced to the computation of power functions $m_k^{\Theta}(m)$ for $(m,k) \in M$ and $m \in \{0,1,\ldots,g_1-1\}$. We remark that this also holds for other characteristics. Moreover, it follows from this lemma, if start point (m,k) of $T(m,k)$ varies in M and if we consider the corresponding power functions then we will at most obtain $s-1$ different power functions. That means further that under the assumptions (3.52) and (3.53) it will not be possible in general to obtain for every pair α and β of given error probabilities an admissible WLRT for

H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ whose power function $M(\theta)$ satisfies

$$M(\theta_o) = \alpha \quad \text{and} \quad M(\theta_1) = 1 - \beta .$$

The following lemma presents a way for the computation of the required power functions $m_k^\theta(m)$.

L e m m a 3.6.2. Suppose that Lemma 3.6.1 holds. Then for $m = g_1-1, g_1-2,\dots,1$ we obtain

$$\vec{m}^\theta(m) = \vec{r}^\theta(m,m+1) + C^\theta(m,m+1)\,\vec{m}^\theta(m+1) \tag{3.64}$$

with

$$\vec{m}^\theta(g_1) = \vec{m}^\theta(0) \tag{3.65}$$

where $\vec{m}^\theta(0)$ is the solution of the system of linear equations

$$(E - C^\theta(0,g_1))\,\vec{m}^\theta(0) = \sum_{n=1}^{g_1} \vec{r}^\theta(0,n). \tag{3.66}$$

P r o o f. Relation (3.66) is a special case of Theorem 3.3.1. Since $(0,k) \simeq (g_1,k+g_o)$ for every $k \in K(0)$ we obtain

$$m_k^\theta(0) = m_{k+g_o}^\theta(g_1) \quad \text{for} \quad k \in K(0).$$

This implies (3.65). Repeated application of the formula of total probability provides recursion formulas (3.64). We refer to the proof of Lemma 3.2.2. ■

To obtain an admissible WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at size (α,β) in case of (3.52) and (3.53) we may now proceed as follows. We complete the definition of power function $M_{z,s}$ of $T_{z,s}$ by $M_{0,s}(\theta) = 0$ and $M_{s,s}(\theta) = 1$, $\theta \in \Theta$, and compute for a preliminary value s the power functions

$$M_{z,s}(\theta_o) \quad \text{and} \quad M_{z,s}(\theta_1) \quad \text{for} \quad z = 1,\dots,s-1$$

according to Lemma 3.6.1 and Lemma 3.6.2. Since these functions are monotonically non-decreasing in z on $\{0,1,\dots,s\}$ integers

$$z' \quad \text{and} \quad z'', \quad 0 < z' \leq s, \quad 0 \leq z'' < s,$$

will exist so that

$$M_{z,s}(\theta_1) \geq 1 - \beta \quad \text{for} \quad z \geq z'$$

and

$$M_{z,s} \leq \alpha \quad \text{for} \quad z \leq z''.$$

If $z' \leq z''$ then every test $T_{z,s}$ with $z' \leq z \leq z''$ is an admissible test for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ at size (α,β). This is emphasized

in Fig.3.1. If $z'' < z'$ then an admissible test $T_{z,s}$ does not exist for the considered value of s. In view of a sample size as small as possible our aim should be to choose s as small as possible. The smallest value of s can be found in case of $z' < z''$ by a successive reduction of s or in case of $z' > z''$ by a successive enlargement of s. A good initial value s for s can be obtained by means of the WALD approximations B^* and A^* for B and A. We obtain

$$s^* = -\operatorname{entier}\left(\frac{g_1}{\gamma_1} \ln \frac{\beta}{1-\alpha}\right) - \operatorname{entier}\left(-\frac{g_1}{\gamma_1} \ln \frac{1-\beta}{\alpha}\right).$$

As a rule, the minimum value for s will be less than s^*.

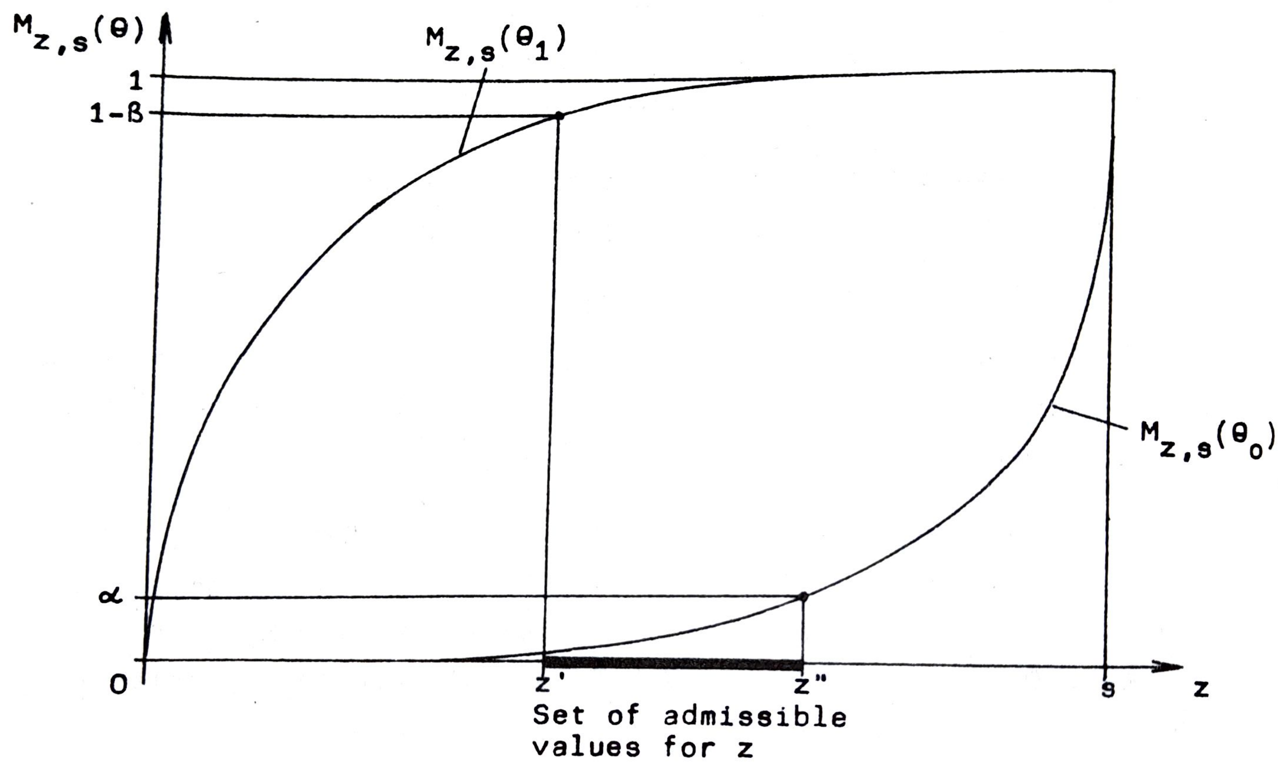

Fig. 3.1. Design of admissible tests

3.7 Grouped observation

Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables with range $\mathfrak{X} \subseteq R^1$ and a distribution indexed by a parameter $\theta \in \Theta$. Suppose that the random variables $\{X_n\}_{n\in\Gamma^+}$ can be observed or are observed only in a restricted manner as follows.

Let $\mathfrak{X}_0, \mathfrak{X}_1, \ldots, \mathfrak{X}_m$, $m \geq 1$, be a given partition of range $\mathfrak{X}$ where

$$\bigcup_{k=0}^{m} \mathfrak{X}_k = \mathfrak{X} \quad \text{and} \quad \mathfrak{X}_i \cap \mathfrak{X}_j = \emptyset \quad \text{for} \quad i \neq j,\ i,j = 0,1,\ldots,m.$$

We assume that only a sequence of random variables $\{X'_n\}_{n\in\Gamma^+}$ is observed where X'_n is defined by

$$X'_n = k \quad \text{if} \quad X_n \in \mathfrak{X}_k, \quad k = 0,1,\dots,m, \; n\in\Gamma^+. \tag{3.67}$$

Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables also the $\{X'_n\}_{n\in\Gamma^+}$ are i.i.d. Hence, for the distribution of $\{X'_n\}_{n\in\Gamma^+}$ we obtain

$$p_\theta(k) = P_\theta(X'_n = k) = P_\theta(X_1 \in \mathfrak{X}_k), \quad k = 0,1,\dots,m, \; n\in\Gamma^+. \tag{3.68}$$

Such an observation scheme is denoted as <u>grouped observation scheme</u>. Certain aspects of theses observation schemes are discussed by [23] and [59].

Now we consider a WLRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on the sequence $\{X'_n\}_{n\in\Gamma^+}$. We suppose that the distribution of X'_1 for $\theta = \theta_1$ is absolutly continuous w.r.t. distribution of X'_1 for $\theta = \theta_0$. Then we obtain for the corresponding likelihood ratios

$$L'_{n,\theta_0,\theta_1} = \prod_{i=1}^{n} \frac{p_{\theta_1}(X'_i)}{p_{\theta_0}(X'_i)}, \quad n\in\Gamma^+,$$

and

$$Z'_{n,\theta_0,\theta_1} = \sum_{i=1}^{n} \ln \frac{p_{\theta_1}(X'_i)}{p_{\theta_0}(X'_i)}, \quad n\in\Gamma^+.$$

Thus, to given stopping bounds B and A, $0 < B < 1 < A < \infty$, we obtain a WLRT $(N', \delta') = \{L'_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ based on $\{X'_n\}_{n\in\Gamma^+}$ where N' and δ' are defined be

$$N' = \begin{cases} \inf\{n \geq 1 : L'_{n,\theta_0,\theta_1} \notin (B,A)\}, & \text{if such an n exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{3.69}$$

and

$$\delta' = \chi_{\{L'_{N,\theta_0,\theta_1} \geq A,\; N' < \infty\}}. \tag{3.70}$$

As a rule, an approximate computation of the power function or the average sample size of (N', δ') by means of the WALD approximations will then only be possible for $\theta = \theta_0$ and $\theta = \theta_1$. The reason is, that the distribution of X'_1 may posess a structure which does no longer allow to apply the approximation methods of Sections 2.1 and 2.7.

In the sequel, we will consider a procedure which will allow us to obtain a WLRT for H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$ by means of an insig-

nificant modification of test variables Z'_{n,θ_o,θ_1}, $n \in \Gamma^+$, so that the modified test possesses approximately the same optimality properties as WLRT (N', δ') defined above but whose characteristics, like e.g. the power function or the moments of the sample size, can be obtained exactly applying the methods of Sections 3.2 to 3.5. An additional advantage of the concept of grouped observation in connection with the subsequent method of the computation of the characteristics will be that a treatment of certain test problems will be possible which can not be solved using the framework of the usual parametric test theory.

We turn again to test (N', δ'). This test can be characterized also as follows. Let c_o and $c_1 > 0$ be given real numbers and let Y_n be a random variable defined by

$$Y_n = \frac{1}{c_1}\left(\ln \frac{p_{\theta_1}(X'_n)}{p_{\theta_o}(X'_n)} + c_o\right), \qquad n \in \Gamma^+. \tag{3.71}$$

Then we have

$$P_\theta(Y_n = \frac{1}{c_1}\left(\ln \frac{p_{\theta_1}(k)}{p_{\theta_o}(k)} + c_o\right)) = P_\theta(X_n \in \mathfrak{X}_k),\ k = 0,1,\dots,m,$$

and

$$Z'_{n,\theta_o,\theta_1} = c_1 \sum_{i=1}^{n} Y_i - c_o\, n\ , \qquad n \in \Gamma^+,$$

so that test (N', δ') can be regarded also as a WLRT for $H_o: \theta = \theta_o$ against $H_1: \theta = \theta_1$ based on sequence $\{Y_n\}_{n\in\Gamma^+}$ where test variables Z'_{n,θ_o,θ_1} have formally the same structure like the test variables Z_{n,θ_o,θ_1} in Section 3.1. Hence, critical inequalities

$$\ln B < Z'_{n,\theta_o,\theta_1} < \ln A, \qquad n \in \Gamma^+,$$

of test (N', δ') can be written as

$$\frac{\ln B}{c_1} + \frac{c_o}{c_1} n < \sum_{i=1}^{n} Y_i < \frac{\ln A}{c_1} + \frac{c_o}{c_1} n, \qquad n \in \Gamma^+,$$

and we obtain

$$N' = \begin{cases} \inf\{n \geq 1: \sum_{i=1}^{n} Y_i \notin \left(\frac{\ln B}{c_1} + \frac{c_o}{c_1} n, \frac{\ln A}{c_1} + \frac{c_o}{c_1} n\right)\}, & \text{if such an } n \text{ exists}, \\ \infty & \text{, otherwise}, \end{cases} \tag{3.72}$$

and

$$\delta' = \chi_{\left\{\sum_{i=1}^{N'} Y_i \geq \frac{\ln A}{c_1} + \frac{c_o}{c_1} N', N' < \infty\right\}}. \tag{3.73}$$

In this form test (N', δ') possesses a structure which may be compared with that of the tests considered in Section 3.1. If, moreover, the random variables $\{Y_n\}_{n \in \Gamma^+}$ are integer-valued random variables we may compute the characteristics of (N', δ') by means of the method introduced in Sections 3.2 to 3.4.
Since c_1 is an arbitrary positive constant we may choose this constant sufficiently small so that random variables $\{Y_n\}_{n \in \Gamma^+}$ can be approximated by integer-valued random variables

$$\bar{Y}_n = \text{entier}(Y_n + 0.5), \qquad n \in \Gamma^+.$$

Then we obtain a test $(\bar{N}, \bar{\delta})$ with

$$\bar{N} = \begin{cases} \inf\left\{n \geq 1 : \sum_{i=1}^{n} \bar{Y}_i \notin \left(\frac{\ln B}{c_1} + \frac{c_o}{c_1} n, \frac{\ln A}{c_1} + \frac{c_o}{c_1} n\right)\right\}, & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{3.74}$$

and

$$\bar{\delta} = \chi_{\left\{\sum_{i=1}^{\bar{N}} Y_i \geq \frac{\ln A}{c_1} + \frac{c_o}{c_1} \bar{N}, \bar{N} < \infty\right\}}. \tag{3.75}$$

This approximation for (N', δ') is the better the smaller c_1 is chosen. Moreover, we may choose integers g_o and $g_1 \neq 0$ so that

$$\frac{c_o}{c_1} \approx \frac{g_o}{g_1}.$$

Hence, test $(\bar{N}, \bar{\delta})$ can be further approximated by test $(\bar{\bar{N}}, \bar{\bar{\delta}})$ defined by

$$\bar{\bar{N}} = \begin{cases} \inf\left\{n \geq 1 : \sum_{i=1}^{n} \bar{Y}_i \notin \left(\frac{g_o \ln B}{g_1 c_o} + \frac{g_o}{g_1} n, \frac{g_o \ln A}{g_1 c_o} + \frac{g_o}{g_1} n\right)\right\}. & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{3.76}$$

and

$$\bar{\bar{\delta}} = \chi_{\left\{\sum_{i=1}^{\bar{\bar{N}}} \bar{Y}_i \geq \frac{g_o \ln A}{g_1 c_o} + \frac{g_o}{g_1} \bar{\bar{N}}, \bar{\bar{N}} < \infty\right\}}. \tag{3.77}$$

The characteristics of this test can be computed by means of the method of Section 3.2. If, for instance, we are interested in the power function of $(\bar{\bar{N}}, \bar{\bar{\delta}})$ we have to solve a system of linear equations of dimension

$$d(m) = \text{card}\left\{k \in \Gamma : \frac{g_o \ln B}{g_1 c_o} + \frac{g_o}{g_1}m < k < \frac{g_o \ln A}{g_1 c_o} + \frac{g_o}{g_1}m\right\}. \qquad (3.78)$$

Again an approximation for d can be obtained by means of the WALD approximations for B and A. To given probabilities α and β, $0 < \alpha, \beta < 1$, $\alpha + \beta < 1$, for an error of first and second kind, respectively, we obtain

$$d(m) \approx d^*(m) = \text{card}\left\{k \in \Gamma : \frac{g_o \ln\frac{\beta}{1-\alpha}}{g_1 c_o} + \frac{g_o}{g_1}m < k < \frac{g_o \ln\frac{1-\beta}{\alpha}}{g_1 c_o} + \frac{g_o}{g_1}m\right\}. \quad (3.79)$$

<u>E x a m p l e 3.7.1</u> The sequential sign test. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having an unknown distribution. For any given $x' \in (-\infty, +\infty)$ parameter θ let be defined by

$$P(X_1 < x') = \theta.$$

Then θ ranges in $[0,1]$. It is desired to discriminate between hypotheses

$$H_o: P(X_1 < x') = \theta_o \quad \text{and} \quad H_1: P(X_1 < x') = \theta_1, \qquad (3.80)$$

$0 < \theta_o < \theta_1 < 1$. Such a problem is called a <u>non-parametric location problem</u>. It may arise, for instance, if a manufacturer is required to have his items exceed a given minimum value - such as life time - without knowing their true distribution. To construct a WLRT for (3.80) we choose

$$\mathfrak{X}_o = [x', \infty) \quad \text{and} \quad \mathfrak{X}_1 = (-\infty, x')$$

and according to (3.67) we obtain

$$X'_n = 0 \quad \text{if} \quad X_n \in [x', \infty)$$

and

$$X'_n = 1 \quad \text{if} \quad X_n \in (-\infty, x'),$$

$n \in \Gamma^+$. For the distribution of X'_n this implies

$$p_\theta(0) = P_\theta(X'_1 = 0) = P(X_1 \in [x', \infty)) = 1 - \theta$$

and

$$p_\theta(1) = P_\theta(X'_1 = 1) = P(X_1 \in (-\infty, x')) = \theta$$

or

$$p_\theta(x) = \theta^x (1-\theta)^{1-x}, \quad x \in \{0,1\}$$

so that the $\{X'_n\}_{n\in\Gamma^+}$ are i.i.d. Bernoulli variables with parameter θ and a WLRT for (3.80) based on $\{X'_n\}_{n\in\Gamma^+}$ can be obtained in precisely the same way as in Example 2.1.0 (i).

A special version of this problem arises if we are interested in testing whether the median of the distribution of X_1 is $\geq x'$ or $< x'$. Then instead of (3.80) we may use the hypotheses

$$H_o: P(X_1 < x') = \theta_o = \frac{1}{2} - \varepsilon \text{ and } H_1: P(X_1 < x') = \theta_1 = \frac{1}{2} + \varepsilon \quad (3.81)$$

for any given ε, $0 < \varepsilon < 1/2$. Then we may imagine that any observation X_n with $X_n \geq x'$ materializes a positive sign, and a negative sign if $X_n < x'$.

Now we shall demonstrate that we will obtain the same test if we proceed as pointed out above. To given c_o and $c_1 > 0$ by (3.71) we obtain

$$Y_n = \frac{1}{c_1}\left[\ln\left(\left(\frac{\theta_1}{\theta_o}\right)^{X_n'}\left(\frac{1-\theta_1}{1-\theta_o}\right)^{1-X_n'}\right) + c_o\right], \quad n \in \Gamma^+.$$

If we choose

$$c_o = -\ln\frac{1-\theta_1}{1-\theta_o} \quad \text{and} \quad c_1 = \ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}$$

we obtain $Y_n = X_n'$, $n \in \Gamma^+$, so that Y_n is an integer-valued random variable and test(N', δ') given by (3.72) and (3.73) is identical with the test obtained by Example 2.1.0 (i). If, moreover,

$$\frac{c_o}{c_1} = -\ln\frac{1-\theta_1}{1-\theta_o}\Bigg/\ln\frac{\theta_1(1-\theta_o)}{\theta_o(1-\theta_1)}$$

is a rational number then we may compute the characteristics of (N', δ') by means of the method of Section 3.2. ■

<u>E x a m p l e 3.7.2</u>. Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. random variables having a normal distribution with unknown mean θ and variance $\sigma^2 = 0.01$. It is desired to discriminate between

$$H_o: \theta = \theta_o = 0 \quad \text{and} \quad H_1: \theta = \theta_1 = 0.1 \qquad (3.82)$$

for a grouped observation scheme characterized by the partition

$$\mathfrak{X}_o = (-\infty, -0.05), \; \mathfrak{X}_1 = [-0.05, 0.05) \text{ and } \mathfrak{X}_2 = [0.05, +\infty)$$

of range $\mathfrak{X} = R^1$ of X_n, $n \in \Gamma^+$. Table 3.7.1 presents the distributions of random variable X_n' given by (3.67) and (3.68) for $\theta = \theta_o$ and $\theta = \theta_1$ as well as the values of likelihood ratio $\ln(p_{\theta_1}(X_n')/p_{\theta_o}(X_n'))$.

k	0	1	2
$p_{\theta_o}(k)$	0.3085	0.3830	0.3085
$p_{\theta_1}(k)$	0.0668	0.2417	0.6915
$\ln(p_{\theta_1}(k)/p_{\theta_o}(k))$	-1.5300	-0.4603	0.8071

Table 3.7.1. Distribution of X_n'

If we choose

$$c_o = 1.5300 \quad \text{and} \quad c_1 = 0.2138$$

then random variable Y_n, $n \in \Gamma^+$, defined by (3.71) takes on the values

$$0, \quad 5.0032 \quad \text{and} \quad 10.9312.$$

This random variable can be approximated by random variable $\bar{Y}_n$, $n \in \Gamma^+$, which takes on the values

$$0, \quad 5 \quad \text{and} \quad 11.$$

Based on the sequence $\{\bar{Y}_n\}_{n \in \Gamma^+}$ we obtain a test $(\bar{N}, \bar{\delta})$ for (3.82) where $\bar{N}$ and $\bar{\delta}$ are defined by (3.74) and (3.75), respectively. In view of the computation of the characteristics of this test ratio c_o/c_1 plays a role. We have

$$\frac{c_o}{c_1} = 7.1529.$$

If we choose $g_o = 7$ and $g_1 = 1$ then we have $\frac{c_o}{c_1} \approx \frac{g_o}{g_1}$ and we may compute the characteristics of test $(\bar{\bar{N}}, \bar{\bar{\delta}})$ for (3.82) where $\bar{\bar{N}}$ and $\bar{\bar{\delta}}$ are defined by (3.76) and (3.77), respectively, applying the method of Section 3.2. To assess the amount of numerical computation we may consider the dimension of vector $\vec{w}^{\theta}(m)$. For instance, if we choose $\alpha = 0.05$, $\beta = 0.05$ and $m = 0$ we obtain $d^*(0) = 27$ by (3.79).■

Further possibilities of an application of this method are tests for a discrimination between two distributions which belong to different parameter families. For instance, we can obtain tests for a discrimination between a normal and Cauchy distribution or a geometrical and a Poisson distribution. Moreover, we may consider hypotheses concerning mixed distributions for which corresponding tests are hardly known. A further advantage of our sequential method in comparison with its fixed sample size counterpart is the fact that it does not require the knowledge of the distribution of test variable $\sum_{i=1}^{n} \bar{Y}_i$ at sample size n but only the distribution of $\bar{Y}_1$ which can be reduced to the distribution of X_1'. Finally, we mention that the method considered above can also be used for an approximate computation of the characteristics of tests based on a sequence of continuous random variables.

3.8 Characteristics of truncated tests

For practical reasons it is often desirable to set a finite upper limit for the number of observations. This leads to truncated sequential tests.

Definition 3.8.1. To a given WLRT $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ a test $(\bar{N}, \bar{\delta})$ is said to be a truncated WLRT if to given $\bar{n}\in\Gamma^+$ and C, $B\leqslant C\leqslant A$, sample size $\bar{N}$ and terminal decision rule $\bar{\delta}$ are defined by

$$\bar{N} = \min\{\bar{n}, N\} \tag{3.83}$$

and

$$\bar{\delta} = \chi_{\{L_{N,\theta_0,\theta_1}\geqslant A,\ N<\bar{n}\}} + \chi_{\{L_{N,\theta_0,\theta_1}\geqslant C,\ N=\bar{n}\}}, \tag{3.84}$$

respectively.

In this context we will also say WLRT $(N, \delta) = \{L_{n,\theta_0,\theta_1}, B, A\}_{n\in\Gamma^+}$ is truncated on stage $\bar{n}$ with the rejection number C. Like in Section 3.2 we will now reduce the computation of the characteristics of $(\bar{N}, \bar{\delta})$ to the computation of the characteristics of a truncated test of type $T(m,k)$ if $\gamma_0/\gamma_1 = g_0/g_1$ holds. For this reason we introduce the following notations.

$$N(m,k;\bar{n}) = \min\{\bar{n}, N(m,k)\}, \tag{3.85}$$

$$Z(m,k;\bar{n}) = g_1\Big(k + \sum_{i=1}^{N(m,k;\bar{n})} X_{m+i}\Big) - g_0(m + N(m,k;\bar{n})), \tag{3.86}$$

$(m,k)\in M$.

Definition 3.8.2. Suppose that (3.14) holds. To given test $T(m,k)$, $(m,k)\in M$, test $(\bar{N}(m,k), \bar{\delta}(m,k))$ is said to be truncated on stage $\bar{n}$, $\bar{n}\in\Gamma_0^+$, with rejection number $\bar{c}$, $(g_1 \ln B)/\gamma_1 \leqslant \bar{c} \leqslant (g_1 \ln A)/\gamma_1$, if sample size $\bar{N}(m,k)$ and terminal decision rule $\bar{\delta}(m,k)$ are given by

$$\bar{N}(m,k) = N(m,k;\bar{n}) \quad \text{and} \tag{3.87}$$

$$\bar{\delta}(m,k) = \chi_{\{Z(m,k;\bar{n})\geqslant \frac{g_1}{\gamma_1}\ln A,\ \bar{N}(m,k)<\bar{n}\}} + \chi_{\{Z(m,k;\bar{n})\geqslant\bar{c},\ \bar{N}(m,k)=\bar{n}\}}, \tag{3.88}$$

respectively. We denote this test by $T(m,k;\bar{n},\bar{c})$.

We will now consider the computation of the characteristics of test $T(m,k;\bar{n},\bar{c})$ which can be represented as expectation value

$$E_\theta w(N(m,k;\bar{n}), Z(m,k;\bar{n})),\ \theta\in\Theta,$$

for any given function $w: \Gamma_0^+\times\Gamma \to R^1$.

Again the characteristic $E_\theta w(\bar{N}, L_{\bar{N},\theta_o,\theta_1})$ of test $(\bar{N}, \bar{\delta})$ can be regarded as a characteristic of a certain $T(m,k:\bar{n},\bar{c})$. If, namely,

$$\frac{\gamma_o}{\gamma_1} = \frac{g_o}{g_1}, \quad g_o, g_1 \in \Gamma^+ \quad \text{and} \quad \bar{c} = \frac{g_1}{\gamma_1} \ln C \tag{3.89}$$

holds, then we have

$$E_\theta w(\bar{N}, L_{\bar{N},\theta_o,\theta_1}) = E_\theta w(N(0,0;\bar{n}), \exp(\frac{\gamma_1}{g_1} Z(0,0;\bar{n})), \quad \theta \in \Theta.$$

In this context we refer to Section 3.2. Then, supposing that the corresponding expectation values exist we have the following lemma.

L e m m a 3.8.1. Suppose that (3.14) holds. Then for every pair $(m,k),(m',k') \in M$ with $(m,k) \cong (m',k')$ we have

$$E_\theta w(N(m,k;\bar{n}), Z(m,k;\bar{n})) = E_\theta w(N(m',k';\bar{n}), Z(m',k';\bar{n})), \quad \theta \in \Theta.$$

P r o o f. This lemma can be proved like Lemma 3.2.1.∎

By means of this lemma we can obtain a recursion formula for the computation of $E_\theta w(N(m,k;\bar{n}), Z(m,k;\bar{n}))$ as follows. To given test $T(m,k;\bar{n},\bar{c})$ and given function $w: \Gamma_o^+ \times \Gamma \to R^1$ we still introduce subsequent notations:

$$w_k^\theta(m;\bar{n}) = E_\theta w(N(m,k;\bar{n}), Z(m,k;\bar{n}))$$

$$\vec{w}^\theta(m;\bar{n}) = \{w_k^\theta(m;\bar{n})\}_{k \in K(m)}$$

$$w_{(1),k}^\theta(m;\bar{n}) = E_\theta w(g_1 + N(m,k;\bar{n}), Z(m,k;\bar{n}))$$

$$\vec{w}_{(1)}^\theta(m;\bar{n}) = \{w_{(1),k}^\theta(m;n)\}_{k \in K(m)}$$

for $(m,k) \in M$, $\bar{n} \in \Gamma_o^+$ and $\theta \in \Theta$. In view of the other notations used in this section we refer to Section 3.2.

L e m m a 3.8.2. Suppose that (3.14) holds. If $\bar{n} > g_1$ then we have

$$\vec{w}^\theta(m;\bar{n}) = \vec{v}^\theta(m) + C^\theta(m, m+g_1)\, \vec{w}_{(1)}^\theta(m;\bar{n} - g_1) \tag{3.90}$$

with

$$\vec{v}^\theta(m) = \sum_{n=1}^{g_1} \sum_{k' \in \bar{K}(m+n)} w(n, g_1 k' - g_o(m+n))\, \vec{c}_{k'}^\theta(m, m+n) \tag{3.91}$$

for every $m \in \Gamma^+$ and $\theta \in \Theta$.

P r o o f. Like in the proof of Lemma 3.2.2 we consider the system of events (3.17) which forms a complete system of pairwise mutually

and exclusive events. Term $\vec{v}^{\theta}(m)$ in (3.90) is obtained like at Lemma 3.2.2 and for sum s_k' given by (3.19) we obtain here

$$s_k' = \sum_{k' \in K(m+g_1)} E_\theta w(N(m,k;\bar{n}),Z(m,k;\bar{n})) \mid C_{kk'}(m,m+g_1)) \cdot \cdot P_\theta(C_{kk'}(m,m+g_1)), \quad k \in K(m).$$

Since $\bar{n} > g_1$ we further obtain

$$P_\theta(C_{kk'}(m,m+g_1)) = c^\theta_{kk'}(m,m+g_1), \quad k \in K(m).$$

The i.i.d.-property of $\{X_n\}_{n\in\Gamma^+}$, $(m,k) \simeq (m+g_1,k')$ for $k' = k+g_o$ and Lemma 3.8.1 imply

$$E_\theta(w(N(m,k;\bar{n}),Z(m,k;\bar{n})) \mid C_{kk'}(m,m+g_1))$$

$$= E_\theta w(g_1 + N(m+g_1,k+g_o;\bar{n}-g_1),Z(m+g_1,k+g_o;\bar{n}-g_1))$$

$$= E_\theta w(g_1 + N(m,k;\bar{n}-g_1),Z(m,k;\bar{n}-g_1))$$

$$= w^\theta_{(1),k}(m;\bar{n}-g_1)$$

for $k \in K(m)$. This completes the proof. ∎

This lemma can be used, for instance, to obtain recursion formulas for the power function or the moments of the sample size of a truncated WLRT. We consider the power function of $T(m,k;\bar{n},\bar{c})$. Denote by $m^\theta_k(m;\bar{n},\bar{c})$ the power function of test $T(m,k;\bar{n},\bar{c})$, $k \in K(m)$, $m \in \Gamma^+$ and $\theta \in \Theta$ where

$$\vec{m}^\theta(m;\bar{n},\bar{c}) = \{m^\theta_k(m;\bar{n},\bar{c})\}_{k\in K(m)}.$$

<u>Lemma 3.8.3</u>. Suppose that (3.14) holds. Then for every $t = 1,2,\ldots$ we have

$$\vec{m}^\theta(m;tg_1,\bar{c}) = \sum_{n=1}^{g_1} \vec{r}^\theta(m,m+n) + C^\theta(m,m+g_1)\, \vec{m}^\theta(m;(t-1)g_1,\bar{c}) \tag{3.92}$$

with the initial condition

$$m^\theta_k(m;0,\bar{c}) = \begin{cases} 1 & \text{for} \quad g_1k - g_om \geq \bar{c} \\ 0 & \text{for} \quad g_1k - g_om < \bar{c} \end{cases} \tag{3.93}$$

where $\vec{r}^\theta(m,m+n)$ is determined by (3.34).

P r o o f. For every $m \in \Gamma^+_o$, $k \in K(m)$ and $\theta \in \Theta$ we have

$$m^\theta_k(m;n,\bar{c}) = E_\theta w(N(m,k;\bar{n}),Z(m,k;\bar{n}))$$

with

$$w(n,z) = \chi_{\{z \geq \bar{c}\}} \quad \text{for} \quad (n,z) \in \Gamma_o^+ \times \Gamma .$$

Applying Lemma 3.8.2 we obtain

$$\vec{m}^{\theta}(m;tg_1,\bar{c}) = \sum_{n=1}^{g_1} \vec{r}^{\theta}(m,m+n) + C^{\theta}(m,m+g_1)\, \vec{w}^{\theta}_{(1)}(m;(t-1)g_1,\bar{c}) \tag{3.94}$$

where $\vec{r}^{\theta}(m,m+n)$ is given by (3.34). Function w above satisfies

$$w(n+g_1,z) = w(n,z) \quad \text{for} \quad (n,z) \in \Gamma_o^+ \times \Gamma .$$

This implies

$$\begin{aligned}\vec{w}^{\theta}_{(1)}(m;(t-1)g_1,\bar{c}) &= \vec{w}^{\theta}(m;(t-1)g_1,\bar{c}) \\ &= \vec{m}^{\theta}(m;(t-1)g_1,\bar{c}).\end{aligned}$$

This, together with (3.94), provides (3.92). The initial condition (3.93) is clear because of $m_k^{\theta}(m;0,\bar{c})$ is per definition the power function of $T(m,k;0,\bar{c})$ which is a test with sample size $N(m,k;0)=0$ ∎

If we are interested in the power function $\bar{M}(\theta)$, $\theta \in \Theta$, of WLRT $(N,\delta) = \{L_{n,\theta_o,\theta_1},B,A\}_{n\in\Gamma^+}$ which is truncated on stage $\bar{n} = tg_1$ with rejection number C, $B \leq C \leq A$, then under the conditions of this lemma we obtain

$$\bar{M}(\theta) = m_0^{\theta}(0,tg_1,\frac{g_1}{\gamma_1}\ln C), \quad \theta \in \Theta .$$

If $\bar{n} = tg_1$ holds for any given $t \in \Gamma_o^+$ our lemma shows further that the corresponding power functions can be obtained directly by the method of iteration. Formally, the computation of the power function $m_k^{\theta}(m;tg_1,\bar{c})$ of test $T(m,k;tg_1,\bar{c})$ according to Lemma 3.8.3 is adequate to the solution of the system of linear equations (3.33) by the method of iteration using the initial condition (3.93). In this sense we can regard $m_k^{\theta}(m;tg_1,\bar{c})$ as the t^{th} approximation for $m_k^{\theta}(m)$, $k \in K(m)$.

Considering test $T(m,k;tg_1,\bar{c})$ we suppose that the upper limit of observations $\bar{n}$ is an integer multiple of g_1. In situations where this assumption is not true we may compute the power function of $T(m,k;\bar{n},\bar{c})$ as follows. Let s be an integer where

$$\bar{n} = tg_1 + s, \quad t \in \Gamma_o^+, \quad 0 \leq s < g_1 .$$

Then in an analogous manner we obtain

$$\vec{m}^{\theta}(m;s,\bar{c}) = \sum_{n=1}^{s-1} \vec{r}^{\theta}(m,m+n) + \vec{\bar{r}}^{\theta}(m,m+s) \tag{3.95}$$

with

$$\vec{r}_k^{\theta}(m,m+s) = \sum_{k'\in K(m+s-1)} c_{kk'}^{\theta}(m,m+n-1)\ P_{\theta}(g_1(k'+X_s)-g_o(m+s) \geq \bar{c})$$

as the probability of acceptance of H_1 by $T(m,k;s,\bar{c})$ on the s^{th} sampling stage. Substituting initial condition (3.93) by (3.95) we may compute $\vec{m}^{\theta}(m;\bar{n},\bar{c})$ according to Lemma 3.8.3 by the method of iteration again.

We now consider the moments of the sample size of a truncated WLRT. Again the computation of theses moments is reduced to the computation of moments of the sample size of certain tests $T(m,k;\bar{n},\bar{c})$. Denote by

$$e_{r,k}^{\theta}(m;\bar{n}) = E_{\theta}N^r(m,k;\bar{n})$$

the r^{th} moment of sample size of test $T(m,k;\bar{n},\bar{c})$, $(m,k)\in M$, $n\in\Gamma_o^+$, $\theta\in\Theta$, $r\in\Gamma_o^+$. Let be $\vec{e}_r^{\theta}(m;\bar{n}) = \{e_{r,k}^{\theta}(m;\bar{n})\}_{n\in\Gamma^+}$. We remark that rejection number $\bar{c}$ does not play any role here.

<u>Lemma 3.8.4</u>. Suppose that (3.14) holds. Then for every $r\in\Gamma_o^+$ and $t = 1,2,\ldots$ we have

$$\begin{aligned}\vec{e}_r^{\theta}(m;tg_1) = &\sum_{n=1}^{g_1} n^r(\vec{a}^{\theta}(m,m+n) + \vec{r}^{\theta}(m,m+n))\\ &+ C^{\theta}(m,m+g_1)\sum_{s=0}^{r-1}\binom{n}{s}g_1^{n-s}\ \vec{e}_s^{\theta}(m;(t-1)g_1)\\ &+ C^{\theta}(m,m+g_1)\ \vec{e}_r^{\theta}(m;(t-1)g_1)\end{aligned} \tag{3.96}$$

with initial conditions

$$\vec{e}_s^{\theta}(m;0) = \begin{cases}\vec{0} & \text{for}\quad s = 1,\ldots,r\\ \vec{1} & \text{for}\quad s = 0\end{cases} \tag{3.97}$$

where $\vec{a}^{\theta}(m,m+n)$ and $\vec{r}^{\theta}(m,m+n)$ are determined by (3.40) and (3,34), respectively.

P r o o f. The initial conditions (3.97) are evident. Applying Lemma 3.8.2 to $w(n,z) = n^r$ we obtain

$$\begin{aligned}\vec{e}_r^{\theta}(m;tg_1) = &\sum_{n=1}^{g_1} n^r(\vec{a}^{\theta}(m,m+n) + \vec{r}^{\theta}(m,m+n))\\ &+ C^{\theta}(m,m+g_1)\ \vec{w}_{(1)}^{\theta}(m;(t-1)g_1)\end{aligned}$$

where

$$w^{\theta}_{(1),k}(m;(t-1)g_1) = E_{\theta}(g_1 + N(m,k;(t-1)g_1))^r$$

$$= \sum_{s=0}^{r} \binom{n}{s} g_1^{n-s} \, E_{\theta}N^{s}(m,k;(t-1)g_1)$$

$$= \sum_{s=0}^{r-1} \binom{n}{s} g_1^{n-s} \, e^{\theta}_{s,k}(m;(t-1)g_1) + e^{\theta}_{r,k}(m;(t-1)g_1)$$

for $k \in K(m)$. This establishes the Lemma.■

If we are again interested in the moments $E_{\theta}N^r$ of sample size of WLRT $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ which is truncated on stage $\bar{n} = tg_1$ then we obtain

$$E_{\theta}N^r = e^{\theta}_{r,0}(0;tg_1), \quad r = 1,2,\ldots,$$

under the conditions of Lemma 3.8.4. If instead of $\bar{n} = tg_1$ we have $\bar{n} = tg_1 + s$, $0 < s < g_1$, $s,t \in \Gamma^+$, we may compute the moments of sample size of test $T(m,k;\bar{n},\bar{c})$ in a similar manner like we have done it for the power function.

We remark that in case of $r = 1$ we have formally the same situation in comparison with the computation of the power function of a truncated test. Under the conditions of Lemma 3.8.4 for $r = 1$ we obtain

$$\vec{e}^{\,\theta}_1(m;tg_1) = \sum_{n=1}^{g_1} n(\vec{a}^{\,\theta}(m,m+n) + \vec{r}^{\,\theta}(m,m+n))$$

$$+ C^{\theta}(m,m+g_1)g_1 \vec{1} + C^{\theta}(m,m+g_1)\, \vec{e}^{\,\theta}_1(m;(t-1)g_1) \tag{3.98}$$

for $t = 1,2,\ldots$ with

$$\vec{e}^{\,\theta}_1(m;0) = \vec{0}. \tag{3.99}$$

Hence, the computation of the average sample size of test $T(m,k;tg_1,\bar{c})$ according to Lemma 3.8.4 is adequate to the solution of the system of linear equations (3.46) by the method of iteration using initial condition (3.99). For $r \geq 2$ we can not interprete relation (3.96) in this sense since the needed terms $\vec{e}^{\,\theta}_s(m;(t-1)g_1)$ depend on t and on each iteration stage we obtain r iteration formulae which are linked together.

3.9 A continuous inspection scheme

We consider a sequence $\{X_n\}_{n\in\Gamma^+}$ of i.i.d. random variables having a distribution depending on parameter $\theta \in \Theta$. We suppose that a parameter change can take place at a random time point $T \in \Gamma^+$ so that the distribution of $X_1,\ldots,X_T$ and $X_{T+1},X_{T+2},\ldots$ is characterized by parameter $\theta_0 \in \Theta$ and $\theta_1 \in \Theta$, $\theta_0 \neq \theta_1$, respectively. Our aim is to detect this parameter change as soon as possible by observing the sequence $\{X_n\}_{n\in\Gamma^+}$. Sampling schemes for this task are called <u>continuous inspection schemes</u> (CIS). Beginning with DODGES's [25] sampling inspection plans a lot of CISs has been created and a certain survey on these sampling schemes has been given by BOWKER [17].
The optimum property of WALD's likelihood ratio test emphasizes to use repeated WLRTs to detect a parameter change. Such an approach has been considered by PAGE [61], [62] and EWAN, KEMP [32]. The problem which arises in this context is the computation of the characteristics of such sampling schemes. Applying the framework of Sections 3.2 to 3.4 we will present an new method for the computation of the moments of the so-called run length of a CIS which is given by a sequence of WLRTs based on discrete random variables. The CIS under consideration is defined as follows.

Let $\{X_n\}_{n\in\Gamma^+}$ be a sequence of i.i.d. integer-valued random variables with a distribution depending on a parameter $\theta \in \Theta$ where at a random time point $T \in \Gamma^+$ a parameter change may occur from $\theta = \theta_0 \in \Theta$ to $\theta = \theta_1 \in \Theta$. To detect this parameter change we consider WLRT $(N,\delta) = \{L_{n,\theta_0,\theta_1},B,A\}_{n\in\Gamma^+}$ where we suppose that

$$Z_{n,\theta_0,\theta_1} = \ln L_{n,\theta_0,\theta_1} = \gamma_1 \sum_{i=1}^{n} X_i - \gamma_0 n,\ n\in\Gamma^+,$$

$\gamma_1 > 0$, holds. Denote by M the continuation region of (N,δ) given by (3.4), then for every $(m,k) \in M$ a CIS can be defined in the following manner.

(i) We start with test $T(m,k)$ and continue sampling according to this test until H_0 or H_1 is accepted.

(ii) If H_1 is accepted by $T(m,k)$ the sampling procedure is terminated and this termination is interpreted as a possible hint to a parameter change from θ_0 to θ_1.

(iii) If H_0 is accepted by $T(m,k)$ on stage $N(m,k) = n'$, $n' \in \Gamma^+$, then test $T(m,k)$ does not provide any indication of a parameter change from θ_0 to θ_1, and we continue sampling by starting a new test $T(m,k_0)$ for H_0 against H_1 based on sequence $\{X_{m+n'+k}\}_{k\in\Gamma^+}$

where k_o is a given integer with

$$\frac{\gamma_0}{\gamma_1} m + \frac{\ln B}{\gamma_1} < k_o < \frac{\gamma_0}{\gamma_1} m + \frac{\ln A}{\gamma_1} .$$

(iv) If H_1 is accepted by $T(m,k_o)$ then the sampling procedure is terminated with a hint to a possible parameter change.
(v) If H_o is accepted by $T(m,k_o)$ on stage $N(m,k_o) = n"$ we start a further test $T(m,k_o)$ for H_o against H_1 based on $\{X_{m+n'+n"+k}\}_{k\in\Gamma^+}$ and proceed like it is formulated by (iv) and (v). This procedure is repeated until hypothesis H_1 is accepted for the first time. We denote such a sampling scheme by $CIS(m,k,k_o)$.

The interesting quantity of this CIS is the number of observations until the first decision for H_1. We shall call it the <u>run length</u> of $CIS(m,k,k_o)$ and denote it by $L_{k_o}(m,k)$. It is evident that for any given $\theta \in \Theta$ for which the probability $m_0^\theta(k_o)$ of acceptance of H_1 by test $T(m,k_o)$ is non-zero we have

$$P_\theta(L_{k_o}(m,k) < \infty) = 1.$$

That means that $CIS(m,k,k_o)$ terminates with probability one whenever $m_0^\theta(k_o) > 0$. If $\theta = \theta_o$ and if a parameter change does not occur the a termination of $CIS(m,k,k_o)$ is non-desired and should happen as rarely as possible. If otherwise θ changes from θ_o to θ_1 a termination should occur as soon as possible. Hence, the most important characteristics assessing the properties of $CIS(m,k,k_o)$ are the moments of run length. We introduce the following notations.

- $K_{k_o}(m,k)$ - random number of tests which are required by $CIS(m,k,k_o)$ until the acceptance of H_1 for the first time;
- $N^{(1)}(m,k)$ - sample size of test $T(m,k)$;
- $N^{(i)}(m,k_o)$ - sample size of i^{th} test in the sequence of tests according to $CIS(m,k,k_o)$, $i = 2,3,\ldots$

Then we have

$$L_{k_o}(m,k) = N^{(1)}(m,k) + \sum_{i=2}^{K_{k_o}(m,k)} N^{(i)}(m,k_o)$$

and the following lemma holds.

<u>Lemma 3.9.1</u>. Consider $CIS(0,k_o,k_o)$. If

$$m_0^\theta(k_o) = P_\theta(\text{Acceptance of } H_1 \text{ by } T(0,k_o)) > 0$$

then we have

(i) $P_\theta(K_{k_o}(0,k_o) = j) = (1 - m^\theta_{k_o}(0))^{j-1}\, m^\theta_{k_o}(0),\ j \in \Gamma^+,$ (3.100)

and

(ii) $$E_\theta L_{k_o}(0,k_o) = \frac{E_\theta N(0,k_o)}{m^\theta_{k_o}(0)}\,. \tag{3.101}$$

P r o o f. Relation (3.100) is an immediate consequence of the definition of CIS$(0,k_o,k_o)$. Hence, the number $K_{k_o}(0,k_o)$ of tests $T(0,k_o)$ until the acceptance of H_1 for the first time is geometrically distributed and we obtain

$$E_\theta K_{k_o}(0,k_o) = 1/m^\theta_{k_o}(0). \tag{3.102}$$

Consider $E_\theta L_{k_o}(0,k_o)$. Then we obtain

$$E_\theta L_{k_o}(0,k_o) = \sum_{j=1}^{\infty} E_\theta(L_{k_o}(0,k_o) \mid K_{k_o}(0,k_o) = j) P_\theta(K_{k_o}(0,k_o)=j). \tag{3.103}$$

Denote by $A^{(i)}$ and $R^{(i)}$ the events of acceptance of H_o and H_1 of i^{th} test $T(0,k_o)$ of CIS$(0,k_o,k_o)$, $i \in \Gamma^+$, respectively. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables we obtain

$$\begin{aligned}
E_\theta(L_{k_o}(0,k_o) \mid K_{k_o}(0,k_o) = j)
&= E_\theta\Big(\sum_{i=1}^{j} N^{(i)}(0,k_o) \mid A^{(1)}\cdots A^{(j-1)}R^{(j)}\Big)\\
&= \sum_{i=1}^{j-1} E_\theta(N^{(i)}(0,k_o) \mid A^{(i)}) + E_\theta(N^{(j)}(0,k_o) \mid R^{(j)})\\
&= (j-1)\, E_\theta(N(0,k_o) \mid A^{(1)}) + E_\theta(N(0,k_o) \mid R^{(1)})
\end{aligned}$$

for every $j \in \Gamma^+$. This, together with (3.103), (3.102) and $m^\theta_{k_o}(0) > 0$, implies

$$\begin{aligned}
E_\theta L_{k_o}(0,k_o) &= E_\theta(N(0,k_o)\mid A^{(1)})\cdot E_\theta K_{k_o}(0,k_o)\\
&\quad - E_\theta(N(0,k_o)\mid A^{(1)}) + E_\theta(N(0,k_o)\mid R^{(1)})
\end{aligned}$$

$$= \frac{E_\theta(N(0,k_o)\mid A^{(1)})}{m^\theta_{k_o}(0)} - E_\theta(N(0,k_o)\mid A^{(1)}) + E_\theta(N(0,k_o)\mid R^{(1)})$$

$$= \frac{1}{m^{\theta}_{k_o}(0)} ((1 - m^{\theta}_{k_o}(0))\, E_{\theta}(N(0,k_o)|A^{(1)}) + m^{\theta}_{k_o}(0)\, E_{\theta}(N(0,k_o)|R^{(1)}))$$

$$= E_{\theta}N(0,k_o)/m^{\theta}_{k_o}(0)$$

and the proof is complete. ■

We remark for a correct interpretation of the average run length that the formulae of Lemma 3.9.1 hold for a fixed parameter $\theta \in \Theta$. If, for instance, a parameter change occurs from θ_o to θ_1 at the time point $T = t$ and if we have reached by CIS$(0,k_o,k_o)$ lattice point $(t,k_t) \in M$ then the average run length $E_{\theta_1} L_{k_o}(0,k_o)$ can be interpreted as follows. If

$$k_t \geq k_o + \frac{\gamma_o}{\gamma_1} t + \frac{\ln B}{\gamma_1}$$

then $E_{\theta_1} L_{k_o}(0,k_o)$ is an upper bound for the average number of observations until termination according to our CIS after the parameter change. Especially, if k_o is chosen by

$$k_o = k^* = \min\left\{k \in \Gamma : k > \frac{\ln B}{\gamma_1}\right\}$$

and if further lattice points do not exist with a smaller distance to the straight line of acceptance of $T(0,k_o)$ than point $(0,k_o)$ then average run length $E_{\theta_1} L_{k^*}(0,k^*)$ is a general upper bound for the average number of observations until sampling termination after the parameter change.

Lemma 3.9.1 shows that the computation of the average run length of CIS$(0,k_o,k_o)$ can be reduced to the computation of average sample number $E_{\theta}N(0,k_o)$ and power function $m^{\theta}_{k_o}(0)$ of test $T(0,k_o)$. If the slope of the straigt line of acceptance of $T(0,k_o)$ is rational then these quantities can be computed by means of the method presented in Sections 3.2 to 3.4. Moreover, for a rational slope it will be possible to compute the moments of the run length in a direct manner reducing this problem to that of solving of a system of linear equations. We shall demonstrate this possibility here only for the average run length.

L e m m a 3.9.2. Consider CIS(m,k,k_o), $(m,k) \in M$, $k_o \in K(m)$. Suppose that

$$\frac{\gamma_o}{\gamma_1} = \frac{g_o}{g_1}, \quad g_o, g_1 \in \Gamma, \quad g_1 > 0.$$

Then we have

$$(E - C^{\theta}(m,m+g_1) - C_o^{\theta})\ \overrightarrow{E_\theta L_{k_o}(m)}$$

$$= \sum_{n=1}^{g_1} n(\vec{a}^{\theta}(m,m+n) + \vec{r}^{\theta}(m,m+n)) + g_1\ C^{\theta}(m,m+g_1)\ \vec{1} \qquad (3.104)$$

with

$$C_o^{\theta} = \left(\sum_{n=1}^{g_1} \vec{a}^{\theta}(m,m+n), \vec{0}, \ldots, \vec{0} \right)$$

and

$$\overrightarrow{E_\theta L_{k_o}(m)} = \left\{ E_\theta L_{k_o}(m,k) \right\}_{k \in K(m)}$$

where $C^{\theta}(m,m+g_1)$ is the transition matrix considered in Lemma 3.2.2 and $\vec{a}^{\theta}(m,m+n)$ and $\vec{r}^{\theta}(m,m+n)$ are given by (3.40) and (3.34), respectively.

P r o o f. Denote by $A_k(m+n)$ and $R_k(m+n)$ for $k \in K(m)$ and $n = 1,\ldots, g_1$ the events of acceptance by $T(m,k)$ on n^{th} sampling stage, respectively. Denote by $C_{kk'}(m,m+g_1)$ for $k \in K(m)$ and $k' \in K(m+g_1)$ the event of reaching lattice point $(m+g_1,k')$ by $T(m,k)$. Then system

$$\left\{ \left\{ A_k(m+n) \right\}_{n=1}^{g_1}, \left\{ R_k(m+n) \right\}_{n=1}^{g_1}, \left\{ C_{kk'}(m,m+g_1) \right\}_{k \in K(m+g_1)} \right\}$$

forms a complete system of pairwise mutually and exclusive events. The corresponding probabilities of these events have been introduced on Sections 3.2 and 3.3. We refer to Lemma 3.2.4, Theorem 3.3.1 and Theorem 3.3.2 and obtain

$$P_\theta(A_k(m+n)) = a_k^{\theta}(m,m+n),\ k \in K(m),\ n = 1,\ldots,g_1,$$

$$P_\theta(R_k(m+n)) = r_k^{\theta}(m,m+n),\ k \in K(m),\ n = 1,\ldots,g_1,$$

$$P_\theta(C_{kk'}(m,m+g_1)) = c_{kk'}^{\theta}(m,m+g_1),\ k \in K(m),\ k' \in K(m+g_1).$$

Applying the formula of total probability we obtain

$$E_\theta L_{k_o}(m,k) = \sum_{n=1}^{g_1} (E_\theta(L_{k_o}(m,k) | A_k(m+n)) a_k^{\theta}(m,m+n)$$

$$+ E_\theta(L_{k_o}(m,k) | R_k(m+n)) r_k^{\theta}(m,m+n))$$

$$+ \sum_{k' \in K(m+g_1)} E_\theta(L_{k_o}(m,k) | C_{kk'}(m,m+g_1)) c_{kk'}^{\theta}(m,m+g_1) \qquad (3.105)$$

for $k \in K(m)$.

In particular, according to the definition of $CIS(m,k,k_o)$ we obtain

$$E_\theta(L_{k_o}(m,k)|A_k(m+n)) = n + E_\theta L^*_{k_o}(m,k_o)$$

for $k \in K(m)$ and $n = 1,\dots,g_1$ where $E_\theta L^*_{k_o}(m,k_o)$ denotes the average run length of $CIS(m,k_o,k_o)$ based on sequence $\{X_{m+n+j}\}_{j\in\Gamma^+}$. Since the $\{X_n\}_{n\in\Gamma^+}$ are assumed to be i.i.d. random variables we have

$$E_\theta L^*_{k_o}(m,k_o) = E_\theta L_{k_o}(m,k_o)$$

and therefore

$$E_\theta(L_{k_o}(m,k)|A_k(m+n)) = n + E_\theta L_{k_o}(m,k_o) \quad \text{for} \quad k\in K(m). \tag{3.106}$$

Further we obtain

$$E_\theta(L_{k_o}(m,k)|R_k(m+n)) = n \quad \text{for} \quad k\in K(m) \text{ and } n = 1,\dots,g_1 \tag{3.107}$$

and, in an analogous manner,

$$E_\theta(L_{k_o}(m,k)|C_{kk'}(m,m+g_1)) = g_1 + E_\theta L_{k_o}(m+g_1,k') \tag{3.108}$$

for $k\in K(m)$ and $k'\in K(m+g_1)$. If we put $k' = h + g_o$, by the i.i.d.-property of $\{X_n\}_{n\in\Gamma^+}$ and $(m,h) \cong (m+g_1,h+g_o)$ for $h\in K(m)$ we obtain

$$E_\theta L_{k_o}(m+g_1,k') = E_\theta L_{k_o}(m+g_1,h+g_o) = E_\theta L_{k_o}(m,h) \tag{3.109}$$

for $h\in K(m)$. Putting together (3.105) to (3.109) we obtain

$$E_\theta L_{k_o}(m,k) = \sum_{n=1}^{g_1} n(a_k^\theta(m,m+n) + r_k^\theta(m,m+n))$$
$$+ E_\theta L_{k_o}(m,k_o)\sum_{n=1}^{g_1} a_k^\theta(m,m+n) + g_1 \sum_{k'\in K(m+g_1)} c_{kk'}^\theta(m,m+g_1)$$
$$+ \sum_{h\in K(m)} E_\theta L_{k_o}(m,h)\, c_{kh+g_o}^\theta(m,m+g_1)$$

for $k\in K(m)$. Writing these equations in matrix form we obtain

$$\overrightarrow{E_\theta L_{k_o}(m)} = \sum_{n=1}^{g_1} n(\vec{a}^\theta(m,m+n) + \vec{r}^\theta(m,m+n))$$
$$+ \left(\sum_{n=1}^{g_1}\vec{a}^\theta(m,m+n),\vec{0},\dots,\vec{0}\right)\overrightarrow{E_\theta L_{k_o}(m)}$$

$$+ g_1 C^\theta(m,m+g_1)\,\vec{1} + C^\theta(m,m+g_1)\,\overrightarrow{E_\theta L_{k_o}(m)}$$

and the lemma is established. ■

Formula (3.104) is quite similar to formula (3.46) for the computation of average sample size $E_\theta N(m,k)$ of $T(m,k)$ for $k \in K(m)$. To compute vector $\overrightarrow{E_\theta L_{k_o}(m)}$ we again only need elementary matrix operations which can be implemented easely on a computer. The computation of the needed quantities for (3.104) can be realized like in Sections 3.2 and 3.3. In a similar manner we may obtain systems of linear equations for higher moments of run length. We present the result for the second moment of run length without proof.

Lemma 3.9.3. Suppose that Lemma 3.9.2 holds. Then we have

$$(E - C^\theta(m,m+g_1) - C_o^\theta)\,\overrightarrow{E_\theta L_{k_o}^2(m)} = \sum_{n=1}^{g_1} n^2(\vec{a}^\theta(m,m+n)+\vec{r}^\theta(m,m+n))$$

$$+ 2\,E_\theta L_{k_o}(m,k_o) \sum_{n=1}^{g_1} n\,\vec{a}^\theta(m,m+n)$$

$$+ g_1^2 C^\theta(m,m+g_1)\,\vec{1} + 2g_1 C(m,m+g_1)\,\overrightarrow{E_\theta L_{k_o}(m)} \quad (3.110)$$

with

$$\overrightarrow{E_\theta L_{k_o}^2(m)} = \left\{E_\theta L_{k_o}^2(m,k)\right\}_{k\in K(m)}$$

where the other quantities are defined like in Lemma 3.9.2.

In view of an application of our CIS the choice of the stopping bounds B and A of the underlying test plays a role. These stopping bounds should be chosen in such a manner that the average run length $E_\theta L_{k_o}(m,k_o)$ is large for $\theta = \theta_o$ and small for $\theta = \theta_1$. We shall say CIS(m,k_o,k_o) is valid if to given bounds l_o and l_1, $0 < l_1 < l_o$, the average run length satisfies inequalities

$$E_{\theta_o} L_{k_o}(m,k_o) \geq l_o \quad \text{and} \quad E_{\theta_1} L_{k_o}(m,k_o) \leq l_1. \quad (3.111)$$

It is not possible to obtain an explicite formula which provides to given bounds l_o and l_1 corresponding values for B and A so that CIS(m,k_o,k_o) is valid in the sense above. Indeed, by means of Lemma 3.9.1 in connection with the corresponding WALD approximations of test $T(m,k_o)$ we may obtain approximations for B and A so that CIS (m,k_o,k_o) is valid in the sense of these approximations.

In doing this we consider the particular case CIS(0,0,0). If (N,δ) is a WLRT for H_o: $\theta = \theta_o$ against H_1: $\theta = \theta_1$ with probabilities α and β of an error of first and second kind, respectively, we obtain

$$E_{\theta_o} L_o(0,0) = \frac{E_{\theta_o} N}{\alpha} \quad \text{and} \quad E_{\theta_1} L_o(0,0) = \frac{E_{\theta_1} N}{1 - \beta} .$$

Hence, α has a much greater influence on average run length than β. Approximating $E_{\theta_o} N$ and $E_{\theta_1} N$ by the corresponding WALD approximations, given by (2.229) and (2.230), respectively, we obtain the approximations

$$E_{\theta_o} L_o(0,0) \approx \frac{(1 - \alpha) \ln \frac{\beta}{1 - \alpha} + \alpha \ln \frac{1 - \beta}{\alpha}}{\alpha E_{\theta_o} Z_{1,\theta_o,\theta_1}} \tag{3.112}$$

and

$$E_{\theta_1} L_o(0,0) \approx \frac{\beta \ln \frac{\beta}{1 - \alpha} + (1 - \beta) \ln \frac{1 - \beta}{\alpha}}{(1 - \beta) E_{\theta_1} Z_{1,\theta_o,\theta_1}} . \tag{3.113}$$

By suitable choice of α and β we may reach that right-hand sides of (3.112) and (3.113) satisfy the validity criterion (3.111). For these values of α and β we may determine the WALD approximations for stopping bounds B and A according to (2.31) by

$$B = \frac{\beta}{1 - \alpha} \quad \text{and} \quad A = \frac{1 - \beta}{\alpha} .$$

These values for B and A can be successively improved by calculating average run lengths $E_{\theta_o} L_o(0,0)$ and $E_{\theta_1} L_o(0,0)$ by means of Lemma 3.9.1 or 3.9.2 and comparing these values with the approximations given by the right-hand sides of (3.112) and (3.113). This comparison provides hints how we have to change α and β to obtain a valid CIS.

4. Discrimination among $k > 2$ hypotheses

There are situations where we may be interested in discrimination between more than two hypotheses. The optimality properties of the WLRT emphasize to use several WLRTs for solving such multiple-decision problem. Based on two WLRTs SOBEL, WALD [72] have considered a corresponding procedure for a discrimination among three simple hypotheses concerning the mean of a normal distribution with known variance. This so-called Sobel-Wald-test is used by GHOSH [35] for discrimination among three hypotheses concerning the parameter of a one-parametric exponential family. In comparison with [72] there are presented improved upper bounds for the average sample size. In view of an improvement of the lower bounds for the average sample size, given in [72] , corresponding results were obtained by SIMONS [71] . Investigations into a direct computation of the average sample size and the power functions of Sobel-Wald-test seem to be unknown so far. Other procedures for a discrimination among more than two hypotheses which do not correspond to Sobel-Wald-test but use sequential tests have been considered by [3],[4],[15],[55],[56], [63],[70],[71] et al.

In this section we will consider a version of Sobel-Wald-test for discrimination between k, $k \geq 2$, simple hypotheses concerning the parameter of a one-dimensional exponential family. Let $\{X_n\}_{n \in \Gamma^+}$ be a sequence of i.i.d. random variables having density

$$f_\theta(x) = h(x) \exp(d(\theta)t(x) - c(\theta)), \quad x \in \mathfrak{X}, \quad \theta \in (\underline{\theta}, \bar{\theta}) \subseteq R^1,$$

where c and d are strictly monotonical in θ on $(\underline{\theta}, \bar{\theta})$, $c(\theta) > 0$ for $\theta \in (\underline{\theta}, \bar{\theta})$. Consider the problem of discrimination among hypotheses

$$H_1: \theta = \theta_1, \; H_2: \theta = \theta_2, \ldots, \; H_k: \theta = \theta_k$$

where

$$\underline{\theta} < \theta_1 < \theta_2 < \cdots < \theta_k < \bar{\theta}. \tag{4.1}$$

Analogously to SOBEL, WALD [72] we consider $k-1$ WLRTs defined as follows. For $j = 1,\ldots,k-1$ let W_j be a WLRT $(N_j, \delta_j) = \{L_{n,\theta_j,\theta_{j+1}}, B_j, A_j\}_{n \in \Gamma^+}$ for $H_0: \theta = \theta_j$ against $H_1: \theta = \theta_{j+1}$ based on stopping bounds B_j and A_j, $0 < B_j < 1 < A_j < \infty$. For the further considerations we still suppose that d is increasing in θ on $(\underline{\theta}, \bar{\theta})$, then critical inequalities

$$B_j < L_{n,\theta_j,\theta_{j+1}} < A_j, \quad n \in \Gamma^+$$

of W_j can be written as

$$h_j^a(n) < \sum_{i=1}^{n} t(X_i) < h_j^r(n),\; n \in \Gamma^+$$

where $h_j^a(n)$ and $h_j^r(n)$ are given by

$$h_j^a(n) = \frac{\ln B_j}{\Delta_j} + \gamma_j n$$

and

$$h_j^r(n) = \frac{\ln A_j}{\Delta_j} + \gamma_j n$$

with

$$\Delta_j = d(\theta_{j+1}) - d(\theta_j)$$

and

$$\gamma_j = (c(\theta_{j+1}) - c(\theta_j))/\Delta_j$$

for $j = 1,\ldots,k-1$ and $n \in \Gamma^+$. Since (4.1) we have

$$\gamma_1 < \gamma_2 < \cdots < \gamma_{k-1}$$

(cf. [35] , Proof of Lemma 5.1). We now suppose that stopping bounds $B_1,\ldots,B_{k-1}$ and $A_1,\ldots,A_{k-1}$ are chosen in such a manner that

$$\frac{\ln B_1}{\Delta_1} \leq \cdots \leq \frac{\ln B_{k-1}}{\Delta_{k-1}} \qquad (4.2)$$

and

$$\frac{\ln A_1}{\Delta_1} \leq \cdots \leq \frac{\ln A_{k-1}}{\Delta_{k-1}} \,. \qquad \begin{matrix}(4.2)\\(4.3)\end{matrix}$$

Then we obtain

$$h_1^a(n) < h_2^a(n) < \cdots < h_{k-1}^a(n) \qquad (4.4)$$

and

$$h_1^r(n) < h_2^r(n) < \cdots < h_{k-1}^r(n) \qquad (4.5)$$

for $n \in \Gamma^+$.

By means of (4.4) and (4.5) we reach a compatibility of $W_1,\ldots,W_{k-1}$ in the sense that, for instance, the acceptance of hypothesis H_j by test W_j implies the acceptance of $H_{j'}$ by test $W_{j'}$, for all $j' > j$. Moreover, these conditions will facilitate the evaluation of the properties of our test.

The procedure under consideration is defined as follows: Suppose that (4.2) and (4.3) hold. Then based on sequence $\{X_n\}_{n \in \Gamma^+}$ we simultaneously realize tests $W_1,\ldots,W_{k-1}$ where at every sampling stage $n = 1,2,\ldots$ we take one of the four decisions:

(i) We accept H_1 if test W_1 accepts H_1.

(ii) We accept H_j if until this stage H_j is accepted by H_j and H_{j-1},

$j = 2,\ldots,k-1$.

(iii) We accept H_k if test W_{k-1} accepts H_k.

(iv) We otherwise continue sampling by observing X_{n+1}.

We denote this procedure by S. Fig. 4.1 illustrates this procedure for $k = 3$ where a possible sampling path is considered which leads to acceptance of H_2 by S.

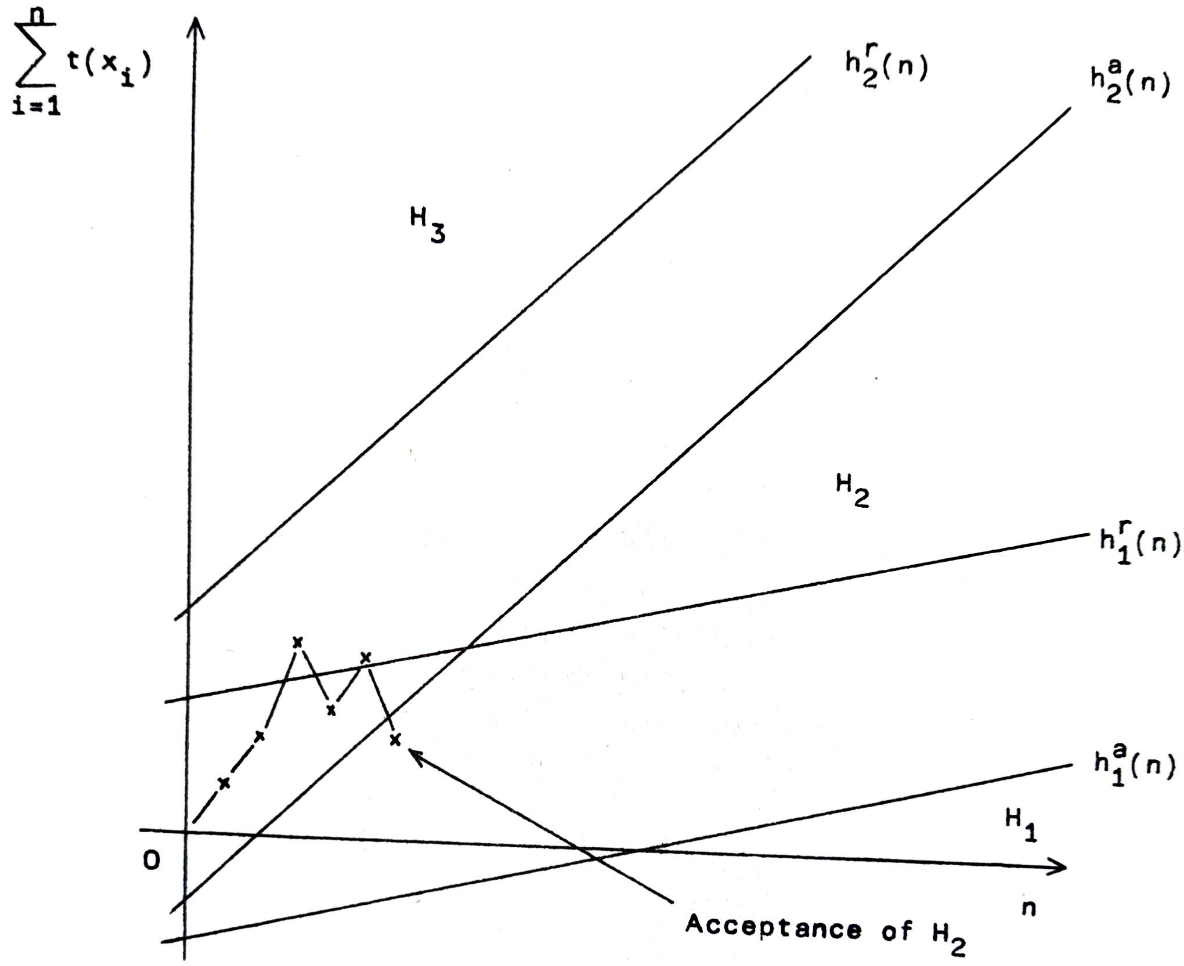

Fig. 4.1. Acceptance and continuation region of S for $k = 3$

Denote by N the sample size of S and let N_j be the sample size of test W_j, $j = 1,\ldots,k-1$. Then, by the definition of S, we have

$$N = \max\left\{N_1,\ldots,N_{k-1}\right\}.$$

This implies

$$N_j \leqslant N < \sum_{i=1}^{k-1} N_i \;, \quad j = 1,\ldots,k-1,$$

and we obtain

$$\max\left\{E_\theta N_1,\ldots,E_\theta N_{k-1}\right\} \leqslant E_\theta N < \sum_{i=1}^{k-1} E_\theta N_i, \quad \theta \in (\underline{\theta},\overline{\theta}).$$

These bounds for $E_\theta N$ are quite useful if the differences are large between neighbouring parameters given by hypotheses $H_1,\dots,H_k$. Further bounds for $E_\theta N$ are considered by [71], [72] and [35]. A direct computation of $E_\theta N$ is a very laborious problem. If S is a procedure based on a sequence of integer-valued random variables where the underlying tests satisfy the assumptions of Sections 3.3 and 3.4 then a direct method of the computation of $E_\theta N$ is presented in [29].

Consider the computation of the power functions of S. Since by S a decision is possible for one of k hypotheses k power functions will be of interest. These power functions can be reduced to the power functions of underlying tests $W_1,\dots,W_{k-1}$. Denote by $M^{(j)}(\theta)$ the probability of acceptance of H_j by S, $j = 1,\dots,k$, $\theta \in (\underline{\theta},\bar{\theta})$, and denote by $M_j(\theta)$ the power function of test W_j, $j = 1,\dots,k-1$, $\theta\ (\underline{\theta},\theta)$.

<u>L e m m a 4.1</u>. For every $\theta \in (\underline{\theta},\bar{\theta})$ the power functions of S satisfy the identities

(i) $\quad M^{(1)}(\theta) = 1 - M_1(\theta)$,

(ii) $\quad M^{(j)}(\theta) = M_{j-1}(\theta) - M_j(\theta)$ for $j = 2,\dots,k-1$,

(iii) $\quad M^{(k)}(\theta) = M_{k-1}(\theta)$.

P r o o f. Denote by A_j and $\bar{A}_j$ the events of the acceptance of hypothesis H_j and H_{j+1} by test W_j, respectively, $j = 1,\dots,k-1$. Then

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_{k-1}$$

and

$$\bar{A}_1 \supseteq \bar{A}_2 \supseteq \cdots \supseteq \bar{A}_{k-1}.$$

Hence, for $\theta \in (\underline{\theta},\bar{\theta})$ we obtain

$$M^{(1)}(\theta) = P_\theta(A_1 \cdots A_{k-1}) = P_\theta(A_1) = 1 - M_1(\theta).$$

For $j = 2,\dots,k-1$ we obtain

$$\begin{aligned} M^{(j)}(\theta) &= P_\theta(\bar{A}_1 \cdots \bar{A}_{j-1} A_j \cdots A_{k-1}) = P_\theta(\bar{A}_{j-1} A_j) \\ &= P_\theta(A_j) - P_\theta(A_{j-1}) = 1 - M_j(\theta) - (1 - M_{j-1}(\theta)) \\ &= M_{j-1}(\theta) - M_j(\theta), \quad \theta \in (\underline{\theta},\bar{\theta}). \end{aligned}$$

Finally, we obtain

$$M^{(k)}(\theta) = P_\theta(\bar{A}_1 \cdots \bar{A}_{k-1}) = P_\theta(\bar{A}_{k-1}) = M_{k-1}(\theta), \quad \theta \in (\underline{\theta},\bar{\theta}).$$

Based on this lemma we may obtain approximations for $M^{(1)}(\theta),\dots, M^{(k)}(\theta)$ if we approximate power functions $M_1(\theta),\dots,M_{k-1}(\theta)$ of $W_1,\dots,W_{k-1}$ by the corresponding WALD approximations. We may obtain exact expressions for $M^{(1)}(\theta),\dots,M^{(k)}(\theta)$ if the method of Section 3.3

can be used to compute $M_1(\theta),\dots,M_{k-1}(\theta)$. Moreover, the identities of Lemma 4.1 can be used to obtain approximations for stopping bounds $B_1,\dots,B_{k-1}$ and $A_1,\dots,A_{k-1}$. For $k = 3$ this problem is discussed by [35].

5. A sequential test for a simultaneous observation of several Bernoulli distributed random variables

Sequential tests for a simultaneous observation of two different Bernoulli distributed random variables on each sampling stage are investigated by [18], [35], [37], [44], [56], [69] and others. Problems of this type may arise in such situations, for instance, where we have to decide between two rival manufacturing processes or medical treatments. These decision problems can be reduced to tests for simple hypotheses. For other aspects we refer to BÜHRINGER et al. [19].

Here we will consider the more general case where we may observe on each sampling stage an m-dimensional, $m \geq 2$, random vector $\vec{X}_n = (X_{n1},\dots,X_{nm})$, $n = 1,2,\dots$, whose components are Bernoulli distributed random variables. Such a situation is given, for instance, if we consider a lot of items where each item is characterized by m several quality characteristics. For the further investigations we generally assume the following:

(i) The random vectors $\vec{X}_1, \vec{X}_2, \dots$ are stochastically independent and identically distributed.

(ii) The components $X_{n1},\dots,X_{nm}$ are stochastically independent Bernoulli distributed random variables with

$$P_{\theta_k}(X_{nk} = x) = \theta_k^x(1-\theta_k)^{1-x} \quad \text{for} \quad x \in \{0,1\},$$

$\theta_k \in [0,1]$, $k = 1,\dots,m$ and $n \in \Gamma^+$. Then, for the distribution of vector $\vec{X}_n$, we obtain

$$P_{\vec{\theta}}(\vec{X}_n = \vec{x}_n) = \prod_{k=1}^{m} \theta_k^{x_k}(1-\theta_k)^{1-x_k}, \quad n \in \Gamma^+,$$

with $\vec{\theta} = (\theta_1,\dots,\theta_m) \in \Theta = \bigtimes_{k=1}^{m} [0,1]$ and $\vec{x}_n \in \bigtimes_{k=1}^{m} \{0,1\}$. That means, that under the above assumptions the distribution of vector $\vec{X}_n$ is completely determined by parameter vector $\vec{\theta}$.

Our aim is to obtain a WLRT for testing hypothesis

$$H_0: \vec{\theta} \in \Theta_0 \quad \text{against} \quad H_1: \vec{\theta} \in \Theta_1 \tag{5.1}$$

where Θ_o and Θ_1 are given disjoint subsets of Θ. We shall later discuss conditions restricting the choice of Θ_o and Θ_1. In testing hypotheses (5.1) we consider a test (N, δ) which is defined as follows. Let $\gamma_o, \gamma_1, \dots, \gamma_m$ be given non-zero real numbers, let Z_n be a random variable defined by

$$Z_n = \sum_{i=1}^{n} \sum_{k=1}^{m} \gamma_k X_{ik} - \gamma_o n, \quad n \in \Gamma^+, \tag{5.2}$$

and let b and a be given stopping bounds, $-\infty < b < 0 < a < +\infty$. Then sample size N and terminal decision rule δ of (N, δ) are defined by

$$N = \begin{cases} \inf\{n \geq 1 : Z_n \notin (b,a)\}, & \text{if such an } n \text{ exists,} \\ \infty & \text{, otherwise,} \end{cases} \tag{5.3}$$

and

$$\delta = \chi_{\{Z_N \geq a, N < \infty\}}, \tag{5.4}$$

respectively. The choice of N and δ in such a manner is motivated by the subsequent example.

<u>Example 5.1</u>. Consider a WLRT for

$$H_o: \vec{\theta} = \vec{\theta}_o \quad \text{against} \quad H_1: \vec{\theta} = \vec{\theta}_1, \tag{5.5}$$

$\vec{\theta}_o, \vec{\theta}_1 \in \Theta$, $\theta_{o,k} \neq \theta_{1,k}$, $k = 1, \dots, m$, based on a sequence $\{\vec{X}_n\}_{n \in \Gamma^+}$ satisfying assumptions (i) and (ii). By the definition of the WLRT we have

$$L_{n,\vec{\theta}_o,\vec{\theta}_1} = \prod_{i=1}^{n} \prod_{k=1}^{m} \left(\left(\frac{\theta_{1,k}}{\theta_{o,k}}\right)^{X_{ik}} \left(\frac{1 - \theta_{1,k}}{1 - \theta_{o,k}}\right)^{1 - X_{ik}} \right)$$

and

$$Z_{n,\vec{\theta}_o,\vec{\theta}_1} = \sum_{i=1}^{n} \sum_{k=1}^{m} \gamma_k X_{ik} - \gamma_o n$$

with

$$\gamma_o = \ln \prod_{k=1}^{m} \left(\frac{1 - \theta_{o,k}}{1 - \theta_{1,k}}\right) \tag{5.6}$$

and

$$\gamma_k = \ln\left(\left(\frac{\theta_{1,k}}{\theta_{o,k}}\right)\left(\frac{1 - \theta_{o,k}}{1 - \theta_{1,k}}\right)\right), \quad k = 1, \dots, m \tag{5.7}$$

$n \in \Gamma^+$. Hence, we obtain test variables $Z_{n,\vec{\theta}_o,\vec{\theta}_1}$, $n \in \Gamma^+$, which have a structure proposed by (5.2). Moreover, if we choose the stopping bounds of our WLRT according to

$$B = \beta \quad \text{and} \quad A = 1/\alpha$$

to given α and β, $0<\alpha,\beta<1$, $\alpha+\beta<1$, then test $(N,\delta) = \{L_{n,\vec{\theta}_0,\vec{\theta}_1}, B, A\}_{n\in\Gamma^+}$ is an admissible test for (5.5) at size (α,β) (cf. Section 2.4). ■

Conversely, a test (N,δ), given by (5.2), (5.3) and (5.4), can be regarded as a WLRT for certain simple hypotheses.

L e m m a 5.1. Consider test (N,δ) given by (5.2), (5.3) and (5.4). For $k = 1,\ldots,m$ let $\theta_{o,k}$ and $\theta_{1,k}$ be defined by

$$\theta_{o,k} = (1-\exp(\gamma_{o,k}))/(1-\exp(\gamma_k)) \tag{5.8}$$

and

$$\theta_{1,k} = \theta_{o,k}\exp(\gamma_k)/\exp(\gamma_{o,k}) \tag{5.9}$$

where

$$\gamma_o = \sum_{k=1}^{m} \gamma_{o,k} \quad \text{and} \quad \gamma_{o,k} \neq 0 \quad \text{for} \quad k = 1,\ldots,m. \tag{5.10}$$

If

$$0<\theta_{o,k}<1, \quad 0<\theta_{1,k}<1 \quad \text{and} \quad \theta_{o,k} \neq \theta_{1,k} \quad \text{for } k = 1,\ldots,m \tag{5.11}$$

then (N,δ) is a WLRT for hypotheses

$$H_o\colon \vec{\theta} = \vec{\theta}_o \quad \text{and} \quad H_1\colon \vec{\theta} = \vec{\theta}_1. \tag{5.12}$$

P r o o f. We have seen in Example 5.1 that under condition (5.11) variables $Z_{n,\vec{\theta}_o,\vec{\theta}_1} = \ln L_{n,\vec{\theta}_o,\vec{\theta}_1}$, $n\in\Gamma^+$, of WLRT $(N,\delta) = \{L_{n,\vec{\theta}_o,\vec{\theta}_1}, B, A\}_{n\in\Gamma^+}$ can be written as

$$Z_{n,\vec{\theta}_o,\vec{\theta}_1} = \sum_{i=1}^{n}\sum_{k=1}^{m} \gamma_k x_{ik} - \gamma_o n, \qquad n\in\Gamma^+,$$

where $\gamma_o, \gamma_1,\ldots,\gamma_m$ are given by (5.6) and (5.7). Conversely, if real numbers $\gamma_o, \gamma_1,\ldots,\gamma_m$ and $\gamma_{o,1},\ldots,\gamma_{o,m}$ are given satisfying (5.10) then vectors $\vec{\theta}_o = (\theta_{o,1},\ldots,\theta_{o,m})$ and $\vec{\theta}_1 = (\theta_{1,1},\ldots,\theta_{1,m})$ can be obtained so that

$$\gamma_{o,k} = \ln((1-\theta_{o,k})/(1-\theta_{1,k})) \tag{5.13}$$

and

$$\gamma_k = \ln\left(\frac{\theta_{1,k}}{\theta_{o,k}}\cdot\frac{1-\theta_{o,k}}{1-\theta_{1,k}}\right) \tag{5.14}$$

for $k = 1,\ldots,m$. Equations (5.13) and (5.14) are equivalent to (5.8) and (5.9). Hence, if (5.11) is true then (N,δ) is a WLRT for (5.12). ■

Since the quantities $\gamma_{o,1},\ldots,\gamma_{m,1}$ used in this lemma are not uniquely determined a continuum of pairs $\vec{\theta}_o,\vec{\theta}_1\in\Theta$ will exist as a rule

so that test (N, δ) given by $\gamma_o, \gamma_1, \ldots, \gamma_m$ is a WLRT for (5.12). To characterize the situations where test (N, δ) given by (5.2), (5.3) and (5.4) can be used for discrimination between two composite hypotheses we consider some gemetrical properties of the WALD approximation of the power function $M(\vec{\theta})$ of (N, δ). Denote here this approximation by $M^*(\vec{\theta})$. If (N, δ) is a WLRT in the sense of Lemma 5.1 we obtain (cf. (2.29))

$$M^*(\vec{\theta}\,') = \begin{cases} \dfrac{1 - \exp(bh)}{\exp(ah) - \exp(bh)}, & \text{if } (\vec{\theta}\,', \vec{\theta}\,'') \overset{h}{\sim} (\vec{\theta}_o, \vec{\theta}_1), \\[2ex] \dfrac{-b}{a - b}, & \text{if } \vec{\theta}\,' \text{ is an exceptional point.} \end{cases}$$

<u>L e m m a 5.2</u>. Consider test (N, δ) given by (5.2), (5.3) and (5.4). Suppose that there exist parameters $\vec{\theta}_o, \vec{\theta}_1 \in \Theta$ so that Lemma 5.1 holds. Denote by Θ^* the set

$$\Theta^* = \{\vec{\theta} \in \Theta : P_{\vec{\theta}}(Z_1 < 0) > 0 \text{ and } P_{\vec{\theta}}(Z_1 > 0) > 0\}.$$

For every $h \in (-\infty, +\infty)$ let Θ^*_h be defined by

$$\Theta^*_h = \{\vec{\theta} \in \Theta^* : \prod_{k=1}^{m} ((\exp(h\gamma_k) - 1)\theta_k + 1) = \exp(h\gamma_o)\} \text{ if } h \neq 0$$

and

$$\Theta^*_h = \{\vec{\theta} \in \Theta^* : \sum_{k=1}^{m} \gamma_k \theta_k - \gamma_o = 0\} \quad \text{if } h = 0.$$

Then we have

$$M^*(\vec{\theta}) = \begin{cases} \dfrac{1 - \exp(bh)}{\exp(ah) - \exp(bh)} & \text{for } \vec{\theta} \in \Theta^*_h,\ h \neq 0, \\[2ex] \dfrac{-b}{a - b} & \text{for } \vec{\theta} \in \Theta^*_o. \end{cases} \tag{5.15}$$

P r o o f. According to Lemma 5.1 test (N, δ) can be regarded as a WLRT for $H_o: \vec{\theta} = \vec{\theta}_o$ against $H_1: \vec{\theta} = \vec{\theta}_1$. Consider the WALD approximation $M^*(\vec{\theta})$ for the power function $M(\vec{\theta})$, $\vec{\theta} \in \Theta^*$, of (N, δ). If $(\vec{\theta}\,', \vec{\theta}\,'') \overset{h}{\sim} (\vec{\theta}_o, \vec{\theta}_1)$ then we obtain

$$M^*(\vec{\theta}\,') = (1 - \exp(bh))/(\exp(ah) - \exp(bh)) \tag{5.16}$$

and $(\vec{\theta}\,', \vec{\theta}\,'') \overset{h}{\sim} (\vec{\theta}_o, \vec{\theta}_1)$ implies

$$L_{1,\vec{\theta}\,',\vec{\theta}\,''} = \prod_{k=1}^{m} \left(\left(\frac{\theta''_k}{\theta'_k} \right)^{X_{1k}} \left(\frac{1 - \theta''_k}{1 - \theta'_k} \right)^{1 - X_{1k}} \right)$$

$$= \left(\prod_{k=1}^{m} \left(\frac{\theta_{1,k}}{\theta_{o,k}}\right)^{X_{1k}} \left(\frac{1-\theta_{1,k}}{1-\theta_{o,k}}\right)^{1-X_{1k}} \right)^{h} = L^{h}_{1,\vec{\theta}_o,\vec{\theta}_1}$$

or

$$\prod_{k=1}^{m} \left(\theta_k''^{X_{1k}} (1-\theta_k'')^{1-X_{1k}} \right) = \exp(hZ_1) \prod_{k=1}^{m} \left(\theta_k'^{X_{1k}} (1-\theta_k')^{1-X_{1k}} \right)$$

where Z_1 is given by (5.2), (5.10), (5.13) and (5.14) for $n = 1$. Hence, we obtain

$$E_{\vec{\theta}'} \prod_{k=1}^{m} \left(\left(\frac{\theta_{1,k}}{\theta_{o,k}}\right)^{X_{1k}} \left(\frac{1-\theta_{1,k}}{1-\theta_{o,k}}\right)^{1-X_{1k}} \right)^{h} = E_{\vec{\theta}'} \exp(hZ_1)$$

$$= E_{\vec{\theta}'} \exp\left(h \left(\sum_{k=1}^{m} \gamma_k X_{1k} - \gamma_o \right) \right)$$

$$= \exp(-h\gamma_o) \prod_{k=1}^{m} \left((\exp(h\gamma_k) - 1)\theta_k' + 1 \right)$$

$$= 1.$$

Otherwise, since $E_{\vec{\theta}'} \exp(hZ_1) = 1$ for $\vec{\theta}' \in \Theta^*_h$ with $h \neq 0$ the term

$$\exp(hZ_1) \prod_{k=1}^{m} \left(\theta_k'^{X_{1k}} (1-\theta_k')^{1-X_{1k}} \right)$$

can be regarded as a density function of $\vec{X}_1$ for a parameter $\vec{\theta}'' \in \Theta$ that satisfies $(\vec{\theta}', \vec{\theta}'') \overset{h}{\sim} (\vec{\theta}_o, \vec{\theta}_1)$. This implies (5.16) for $\vec{\theta}' \in \Theta^*_h$ with $h \neq 0$. Next, it is a well-known fact, see e.g. GHOSH [35], Lemma 3.4, that for $\vec{\theta}' \in \Theta^*$ equation $E_{\vec{\theta}'} \exp(hZ_1) = 1$ has only solution $h = 0$ if $E_{\vec{\theta}'} Z_1 = 0$. That means, that in case of $E_{\vec{\theta}'} Z_1 = 0$ there does not exist a parameter $\vec{\theta}'' \in \Theta$ with $(\vec{\theta}', \vec{\theta}'') \overset{h}{\sim} (\vec{\theta}_o, \vec{\theta}_1)$ so that $\vec{\theta}'$ is an exceptional point. Hence, we obtain

$$M^*(\vec{\theta}') = -b/(a-b) \quad \text{for } E_{\theta'} Z_1 = 0 \text{ and } \theta' \in \Theta^*_o.$$

This completes the proof.■

By this lemma the hypersurfaces are characterized on which $M^*(\vec{\theta})$ is constant. These hypersurfaces are given by the equations

$$\prod_{k=1}^{m} \left((\exp(h\gamma_k) - 1)\theta_k + 1 \right) = \exp(h\gamma_o) \quad \text{for} \quad h \neq 0 \qquad (5.17)$$

and

$$\sum_{k=1}^{m} \gamma_k \theta_k - \gamma_o = 0 \qquad \text{for} \qquad h = 0,$$

respectively.

We discuss some properties of these hypersurfaces.

(i) Let $h \neq 0$. Then we can write (5.17) as

$$\prod_{k=1}^{m} \left(\theta_k - \frac{1}{\exp(h\gamma_k) - 1}\right) = \exp(h\gamma_o)\left(\prod_{k=1}^{m} (\exp(h\gamma_k) - 1)\right)^{-1}.$$

This equation is for the particular case $m = 2$ the algebraic equation of a hyperbola with asymptotes

$$\theta_k = -1/(\exp(h\gamma_k) - 1) \qquad \text{for} \quad k = 1,2,$$

see also MALY [56] .

(ii) If we may suppose that $\theta_i\theta_j \approx 0$ for $i,j = 1,\ldots,m$ then we have

$$\exp(-h\gamma_o)\prod_{k=1}^{m}((\exp(h\gamma_k) - 1)\theta_k + 1)$$

$$\approx \exp(-h\gamma_o)\left(\sum_{k=1}^{m}((\exp(h\gamma_k) - 1)\theta_k) + 1\right).$$

Thus, we may approximate the hypersurface (5.17) for sufficient small values of $\theta_1,\ldots,\theta_m$ by a hyperplane with algebraic equation

$$\sum_{k=1}^{m} (\exp(h\gamma_k) - 1)\theta_k/(\exp(h\gamma_o) - 1) = 1.$$

This may be of interest in certain quality control problems.

(iii) We may consider relation (5.17) as an implicitly given function with dependent variable θ_k, $k \in \{1,\ldots,m\}$. Then, by (5.17) we have

$$\theta_k = \exp(h\gamma_o)\left(1\Big/\prod_{\substack{j=1\\ j\neq k}}^{m}((\exp(h\gamma_j) - 1)\theta_j + 1) - 1\right).$$

If $\gamma_k > 0$ for $k = 1,\ldots,m$ one may show that these functions are convex for $h > 0$ and concave for $h < 0$, respectively.

We now consider some monotonicity properties of $M^*(\vec{\theta})$. In case of $m = 1$ we know that $M^*(\theta)$ is strictly monotonical in θ on Θ. For an appropriate assertion in the m-dimensional case, $m > 1$, we have to define in which sense we want to compare several parameters of Θ. We shall say parameter $\vec{\theta}' \in \Theta$ is <u>better</u> than parameter $\vec{\theta}'' \in \Theta$

(write: $\vec{\theta}' \prec \vec{\theta}''$) iff $\theta_k' \le \theta_k''$ for $k = 1,\dots,m$ and there at least exists one $k^* \in \{1,\dots,m\}$ with $\theta_{k^*}' < \theta_{k^*}''$.

<u>L e m m a 5.3</u>. Suppose that Lemma 5.2 holds. If $\gamma_k > 0$ for $k = 1,\dots,m$ then we have

$$M^*(\vec{\theta}') < M^*(\theta'') \quad \text{for} \quad \vec{\theta}' \prec \vec{\theta}'', \quad \vec{\theta}', \vec{\theta}'' \in \Theta^*. \tag{5.18}$$

P r o o f. Without any loss of generality we may suppose that $\vec{\theta}'$ and $\vec{\theta}''$ differ only in one component, say $\theta_{k^*}' < \theta_{k^*}''$. To given $\vec{\theta} \in \Theta^*$ we consider function

$$\varphi_{\vec{\theta}}(h) = E_{\vec{\theta}} \exp(hZ_1) = \exp(-h\gamma_0) \prod_{k=1}^{m} ((\exp(h\gamma_k) - 1)\theta_k + 1) \tag{5.19}$$

as a function of h, $-\infty < h < +\infty$, which is convex from below with $\lim_{h \to \pm\infty} \varphi_\theta(h) = \infty$ for $\vec{\theta} \in \Theta^*$ (see e.g. [35], Proof of Lemma 3.4).
Denote by $h(\vec{\theta})$ for $\vec{\theta} \in \Theta^*$ the uniquely determined non-zero solution of $\varphi_{\vec{\theta}}(h) = 1$ if $E_{\vec{\theta}} Z_1 \neq 0$. Let $h(\vec{\theta})$ be zero if $E_{\vec{\theta}} Z_1 = 0$ (we refer again to [35], Lemma 3.4) then the following cases may arise:
(i) $E_{\vec{\theta}'} Z_1 < 0$: Then we have $h(\vec{\theta}) > 0$ and (5.19) implies $\exp(h\gamma_k) - 1 > 0$ for $h > 0$ and $\gamma_k > 0$, $k = 1,\dots,m$. The convexity of $\varphi_{\vec{\theta}'}(h)$ implies

$$\begin{aligned} 1 = \varphi_{\vec{\theta}'}(h(\vec{\theta}')) &= \exp(-h(\vec{\theta}')\gamma_0) \prod_{k=1}^{m} ((\exp(h(\vec{\theta}')\gamma_k) - 1)\theta_k' + 1) \\ &< \exp(-h(\vec{\theta}')\gamma_0) \prod_{k=1}^{m} ((\exp(h(\vec{\theta}')\gamma_k) - 1)\theta_k'' + 1) \\ &= \varphi_{\vec{\theta}''}(h(\vec{\theta}')). \end{aligned}$$

Hence, we obtain $h(\vec{\theta}') > h(\vec{\theta}'')$.
(ii) $E_{\vec{\theta}'} Z_1 = 0$: Then we have $h(\vec{\theta}') = 0$ and by $\gamma_k > 0$ for $k = 1,\dots,m$ and $\theta_{k^*}' < \theta_{k^*}''$ we obtain

$$0 = E_{\vec{\theta}'} Z_1 = \sum_{k=1}^{m} \gamma_k \theta_k' - \gamma_0 < \sum_{k=1}^{m} \gamma_k \theta_k'' - \gamma_0 = E_{\vec{\theta}''} Z_1 .$$

For $E_{\vec{\theta}''} Z_1 > 0$ we obtain $h(\vec{\theta}'') < 0$. This implies $h(\vec{\theta}') > h(\vec{\theta}'')$.
(iii) $E_{\vec{\theta}'} Z_1 > 0$: Then we have $h(\vec{\theta}') < 0$ and (5.19), $\exp(h\gamma_k) - 1 < 0$ for $h < 0$ and $\gamma_k > 0$ for $k = 1,\dots,m$ and the convexity of $\varphi_{\vec{\theta}'}(h)$ imply

$$1 = \varphi_{\vec{\theta}'}(h(\vec{\theta}')) < \varphi_{\vec{\theta}''}(h(\vec{\theta}')).$$

This provides again $h(\vec{\theta}') > h(\vec{\theta}'')$.
Hence we obtain

$$h(\vec{\theta}') > h(\vec{\theta}'') \quad \text{for} \quad \vec{\theta}' < \vec{\theta}'', \quad \vec{\theta}', \vec{\theta}'' \in \Theta^*. \tag{5.20}$$

Now, if $h(\vec{\theta}') \neq 0$ we have $(\vec{\theta}', \vec{\theta}'') \overset{h(\vec{\theta}')}{\sim} (\vec{\theta}_o, \vec{\theta}_1)$ for $\vec{\theta}' \in \Theta^*$ where $\vec{\theta}_o$ and $\vec{\theta}_1$ are determined like in Lemma 5.1. Thus, we obtain

$$M^*(\vec{\theta}') = (1 - \exp(bh(\vec{\theta}'))/(\exp(ah(\vec{\theta}') - \exp(bh(\vec{\theta}'')). \tag{5.21}$$

If $h(\vec{\theta}') = 0$ then $\vec{\theta}'$ is an exceptional point and we obtain

$$M^*(\vec{\theta}') = - b/(a - b). \tag{5.22}$$

Hence, by (5.21) and (5.22) we obtain

$$M^*(\vec{\theta}') < M^*(\vec{\theta}'') \quad \text{for} \quad h(\vec{\theta}') > h(\vec{\theta}'').$$

This, together with (5.20), provides (2.18).■

The monotonicity properties of $M^*(\vec{\theta})$ can be used to obtain a corresponding admissibility assertion for composite hypotheses.

<u>L e m m a 5.4</u>. We suppose that Lemma 5.2 holds where $\gamma_k > 0$ for $k = 1,\ldots,m$. Let Θ_o and Θ_1 be disjoint subsets of Θ^*. We suppose that finite subsets Θ' and Θ'' of Θ^* exist so that

(i) for every $\vec{\theta} \in \Theta_o$ there exists a $\vec{\theta}' \in \Theta'$ with $\vec{\theta} < \vec{\theta}'$,
(ii) for every $\vec{\theta} \in \Theta_1$ there exists a $\vec{\theta}'' \in \Theta''$ with $\vec{\theta}'' < \vec{\theta}$,
(iii) $E_{\vec{\theta}} Z_1 < 0$ for $\vec{\theta} \in \Theta'$,
(iv) $E_{\theta} Z_1 > 0$ for $\vec{\theta} \in \Theta''$.

Then, by a choice of sufficiently large stopping bounds $|b|$ and a, we may reach

$$M^*(\vec{\theta}) \leq \alpha \quad \text{for } \vec{\theta} \in \Theta_o \tag{5.23}$$

and

$$M^*(\vec{\theta}) \geq 1 - \beta \quad \text{for } \vec{\theta} \in \Theta_1 \tag{5.24}$$

to given α, β, $0 < \alpha, \beta < 1$, $\alpha + \beta < 1$.

P r o o f. For every $\vec{\theta} \in \Theta_o$ there exists a $\vec{\theta}' \in \Theta'$ with $\vec{\theta} < \vec{\theta}'$. By Lemma 5.3 we obtain $M^*(\vec{\theta}) < M^*(\vec{\theta}')$. Since Θ' is a finite set this implies

$$M^*(\vec{\theta}) \leq \max_{\theta' \in \Theta'} M^*(\vec{\theta}') \quad \text{for} \quad \theta \in \Theta_o. \tag{5.25}$$

Denote by $h(\vec{\theta})$ the non-zero solution of $\varphi_{\vec{\theta}}(h) = E_{\vec{\theta}} \exp(hZ_1) = 1$ for $E_{\vec{\theta}} Z_1 \neq 0$, $\vec{\theta} \in \Theta^*$. Then, for $\vec{\theta}' \in \Theta'$ we have $E_{\vec{\theta}'} Z_1 < 0$ and therefore $h(\vec{\theta}') > 0$. Since Θ' is finite this provides

$$\min_{\theta' \in \Theta'} h(\vec{\theta}') = h(\vec{\theta}^*) > 0, \quad \vec{\theta}^* \in \Theta',$$

and, together with (5.25), we obtain

$$M^*(\vec{\theta}) \leq M^*(\vec{\theta}^*) = (1 - \exp(bh(\vec{\theta}^*))/(\exp(ah(\vec{\theta}^*) - \exp(bh(\vec{\theta}^*)), \quad \vec{\theta} \in \Theta_o. \tag{5.26}$$

Analogously, we may obtain a $\vec{\theta}^{**} \in \Theta''$ so that $h(\vec{\theta}^{**}) < 0$ and

$$M^*(\vec{\theta}) \geq M^*(\vec{\theta}^{**})$$

$$= (1 - \exp(bh(\vec{\theta}^{**})))/(\exp(ah(\vec{\theta}^{**}) - \exp(bh(\vec{\theta}^{**})), \quad \vec{\theta} \in \Theta_1, \tag{5.27}$$

This, together with (5.26), provides (5.23) and (5.24) for sufficiently large values of $|b|$ and a. ■

We notice that the conditions (i) to (iv) of this lemma are fulfilled, for instance, if sets Θ_o and Θ_1 can be strongly separated by hyperplane $E_{\vec{\theta}} Z_1 = 0$.

We consider a special case. Let Θ_o and Θ_1 be given by

$$\Theta_o = \{\vec{\theta} \in \Theta : \sum_{k=1}^{m} \frac{\theta_k}{\theta_k^{(\alpha)}} \leq 1\} \tag{5.28}$$

and

$$\Theta_1 = \{\vec{\theta} \in \Theta : \sum_{k=1}^{m} \frac{\theta_k}{\theta_k^{(\beta)}} \geq 1\} \tag{5.29}$$

where $\theta_k^{(\alpha)}$ and $\theta_k^{(\beta)}$ denote given real numbers, $0 < \theta_k^{(\alpha)} < \theta_k^{(\beta)} < 1$ for $k = 1,\dots,m$. Our aim is to discriminate between hypotheses

$$H_o: \vec{\theta} \in \Theta_o \quad \text{against} \quad H_1: \vec{\theta} \in \Theta_1. \tag{5.30}$$

Problems of this type may arise in quality control if we consider a lot of items where the quality of each item is characterized by m several attributive and stochastically independent quality characteristics. Then, the quantity $\theta_k^{(\alpha)}$ may be regarded as the 'conditional' acceptable quality level to the k^{th} quality characteristic if defects may only occur at this quality characteristic. In an analogous manner we may regard $\theta_k^{(\beta)}$ as 'conditional' limiting quality. The following lemma presents a condition for the choice of quantities $\gamma_o, \gamma_1, \dots, \gamma_m$ of test variable Z_1 of test (N,δ) given by (5.2), (5.3) and (5.4).

L e m m a 5.5. Let Θ_o and Θ_1 are defined by (5.28) and (5.29). Suppose that

$$\gamma_k > \gamma_o > 0 \quad \text{for} \quad k = 1,\dots,m \tag{5.31}$$

and

$$\theta_k^{(\alpha)} < \frac{\gamma_o}{\gamma_k} < \theta_k^{(\beta)} \quad \text{for} \quad k = 1,\dots,m. \tag{5.32}$$

Then we have

$$M^*(\vec{\theta}') < M^*(\vec{\theta}'') \quad \text{for} \quad \vec{\theta}' < \vec{\theta}'' \quad \text{and} \quad \vec{\theta}', \vec{\theta}'' \in \Theta^*.$$

Moreover, for sufficiently large values of $|b|$ and a we may reach

$$M^*(\vec{\theta}) \leq \alpha \qquad \text{for} \quad \vec{\theta} \in \Theta_0 \cap \Theta^* \tag{5.33}$$

and

$$M^*(\vec{\theta}) \geq 1 - \beta \quad \text{for} \quad \vec{\theta} \in \Theta_1 \cap \Theta^* \tag{5.34}$$

to given α, β, $0<\alpha,\beta<1$, $\alpha+\beta<1$.

P r o o f. Since $\gamma_k > 0$ for $k = 1,\ldots,m$ the monotonicity of $M^*(\vec{\theta})$ is an immediate conclusion of Lemma 5.3. Consider the hyperplane $E_{\vec{\theta}} Z_1 = 0$. This hyperplane intersects the θ_k-axis at $\theta_k = \gamma_0/\gamma_k$ for $k = 1,\ldots,m$, and under condition (5.32) this hyperplane strongly separates sets Θ_0 and Θ_1. Applying Lemma 5.4 we obtain (5.33) and (5.34). ■

According to (5.31) and (5.32) the range for the choice of the quantities $\gamma_0, \gamma_1, \ldots, \gamma_m$ is very wide and, by a proper choice of these quantities, we may take influence on the average sample size of our test. Without to discuss this problem in the details a suitable choice seems to be the following if $\theta_k^{(\alpha)}$ and $\theta_k^{(\beta)}$ are sufficiently small for $k = 1,\ldots,m$.

Let $\hat{\theta}_k^{(\alpha)}$ and $\hat{\theta}_k^{(\beta)}$ are for $k = 1,\ldots,m$ real numbers so that to given $\theta_k^{(\alpha)}$ and $\theta_k^{(\beta)}$, $0 < \theta_k^{(\alpha)} < \theta_k^{(\beta)} < 1$,

$$\theta_k^{(\alpha)} \leq \hat{\theta}_k^{(\alpha)} < \hat{\theta}_k^{(\beta)} \leq \theta_k^{(\beta)}, \qquad k = 1,\ldots,m,$$

and

$$\frac{1 - \hat{\theta}_1^{(\alpha)}}{1 - \hat{\theta}_1^{(\beta)}} = \cdots = \frac{1 - \hat{\theta}_m^{(\alpha)}}{1 - \hat{\theta}_m^{(\beta)}} .$$

If we choose

$$\gamma_0 = \ln\left((1 - \hat{\theta}_1^{(\alpha)})/(1 - \hat{\theta}_1^{(\beta)})\right)$$

and

$$\gamma_k = \gamma_0 + \ln\left(\hat{\theta}_k^{(\beta)}/\hat{\theta}_k^{(\alpha)}\right) \qquad \text{for} \quad k = 1,\ldots,m$$

then assumptions (5.31) and (5.32) of Lemma 5.5 are fulfilled and test (N, δ) is admissible in the sense of Lemma 5.5.

We remark that in case of $\vec{\theta} \notin \Theta^*$ we have either

$$\vec{\theta} \in \hat{\Theta} = \{\vec{\theta} \in \Theta : P_{\vec{\theta}}(Z_1 > 0) = 0 \text{ and } P_{\theta}(Z_1 < 0) > 0\}$$

or

$$\vec{\theta} \in \hat{\hat{\Theta}} = \{\vec{\theta} \in \Theta : P_{\vec{\theta}}(Z_1 > 0) > 0 \text{ and } P_{\theta}(Z_1 < 0) = 0\}.$$

The case $P_{\vec{\theta}}(Z_1 = 0) = 1$ is impossible since the $X_{11},\ldots,X_{1m}$ are assumed to be independent. Evidently, we obtain

$$M(\vec{\theta}) = 0 \quad \text{for} \quad \vec{\theta} \in \hat{\Theta} \quad \text{and} \quad M(\vec{\theta}) = 1 \quad \text{for} \quad \vec{\theta} \in \hat{\hat{\Theta}}.$$

Under the conditions of Lemma 5.5 we furthermore obtain

$$P_{\vec{\theta}}(Z_1 > 0) = P_{\vec{\theta}}\left(\bigcup_{k=1}^{m}\{X_{1k} = 1\}\right) = 1 - \prod_{k=1}^{m}(1 - \theta_k),$$

$$P_{\vec{\theta}}(Z_1 < 0) = P_{\vec{\theta}}\left(\bigcap_{k=1}^{m}\{X_{1k} = 0\}\right) = \prod_{k=1}^{m}(1 - \theta_k).$$

Hence we obtain

$$\hat{\Theta} = \left\{\vec{\theta} \in \Theta : \prod_{k=1}^{m}(1 - \theta_k) = 1\right\} = \{\vec{0}\},$$

$$\hat{\hat{\Theta}} = \left\{\vec{\theta} \in \Theta : \prod_{k=1}^{m}(1 - \theta_k) = 0\right\}.$$

That means that set Θ^* contains at least all interior points of Θ.

References

[1] ABRAHAM, J.K. (1969). The local power of sequential tests subject to an expected sample size restriction. Techn. Rep. Department of Statistics, Stanford University.

[2] ALBERT, G.E. (1956). Accurate sequential tests on the mean of an exponential distribution. Ann. Math. Statist. 27, 460-470.

[3] ARMITAGE, P. (1947). Some sequential tests of 'Student's hypothesis. Suppl. J. R. Statist. Soc. 9, 250-263.

[4] ARMITAGE, P. (1950). Sequential analysis with more than two hypotheses, and its relation to discriminant function analysis. JRSS, Ser. B, 12, 137-144.

[5] AROIAN, L.A. (1968). Sequential analysis, direct method. Technometrics 10, 125-132.

[6] AROIAN, L.A., ROBINSON, D.E. (1969). Direct methods for exact truncated sequential tests of the mean of a normal distribution. Technometrics 11, 661-675.

[7] ARROW,K., BLACKWELL, D., GIRSHICK, M.A. (1949). Bayes and minimax solutions of sequential decision problems. Econometrica 17, 213-244.

[8] BAHADUR, R.R. (1954). Sufficiency and statistical decision functions. Ann. Math. Statist. 25, 423-462.

[9] BARRACLOUGH, E.D., PAGE, E.S. (1959). Tables for Walds test for the mean of a normal distribution. Biometrika 46, 169-177.

[10] BARTKY, W. (1943). Multiple sampling with constant probability. Ann. Math. Statist. 14, 363-377.

[11] BAUER, H. (1968). Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. Walter de Gruyter&Co., Berlin.

[12] BERK, R.H. (1970). Stopping times of SPRTs based on exchangeable models. Ann. Math. Statist. 41, 979-990.

[13] BERK, R.H. (1975). Locally most powerful sequential tests. Ann. Statist. 3, 373-381.

[14] BERNSTEIN, P. (1964). Gruppierte Sequenzteste zur Prüfung des Mittelwertes einer 0-1-verteilten Zufallsvariablen. Abhandl. Dtsch. Akad. Wiss. Bln., Kl. Math., Phys. u. Techn. 4, 31-33.

[15] BILLARD, L., VAGHOLKAR, M.K. (1969). A sequential procedure for testing a null hypothesis against a two-sided alternative hypothesis. JRSS, Ser. B, 31,285-294.

[16] BLASBALG, H. (1957). Transformation of the fundamental relationships in sequential analysis. Ann. Math. Statist. 28, 1024-1028.

[17] BOWKER, A.H. (1956). Continuous sampling plans. Proceedings of the Third Berkely Symposium on Mathematical Statistics and Probability, Vol. V. Berkely and Los Angeles, 75-85.

[18] BROSS, I. (1952). Sequential medical plans. Biometrics 8, 188-205.

[19] BÜRINGER, H., MARTIN, H., SCHRIEVER, K.-H. (1980). Nonparametric sequential selection procedures. Birkhäuser Boston.

[20] BURKHOLDER, D.L., WIJSMAN, R.A. (1963). Optimum properties and admissibility of sequential tests. Ann. Math. Statist. 34, 1-18.

[21] BURMAN, J.R. (1946). Sequential sampling formulae for a binomial population. JRSS, Ser. B, 8, 98-103.

[22] CHOW, Y.S., ROBBINS, H., SIEGMUND, D. (1971). Great expectations: The theory of optimal stopping. Houghton Mifflin Company, New York.

[23] COX, D.R. (1957). Note on grouping. JASA 52, 543-547.

[24] DE GROOT, M.H. (1959). Unbiased sequential estimation for binomial populations. Ann. Math. Statist. 30, 80-101.

[25] DODGE, H.F. (1943). A sampling inspection plan for continuous production. Ann. Math. Statist. 14, 264-279.

[26] DÖHLER, R. (1979). Dominierbarkeit und Suffizienz in der Sequentialanalyse. Wiss. Schriftenreihe der Techn. Hochschule Karl-Marx-Stadt, Heft 2/1979, 4-48.

[27] EGER, K.-H. (1977). Zur Berechnung der Operationscharacteristik und des mittleren Stichprobenumfangs bei Mehrfachstichprobenplänen zur Attributprüfung. Wiss. Zeitschr. d. Techn. Hochsch. Karl-Marx-Stadt, 19, 467-475.

[28] EGER, K.-H. (1980). A direct method of the computation of the OC and of the moments of the sample number for SPRTs in the case of discrete random variables. Math. Operationsforsch. Statist., Series Statistics, 11, 499-514.

[29] EGER, K.-H. (1981). Eine direkte Methode zur Berechnung der Charakteristiken des SOBEL-WALD-Tests zur Entscheidung zwischen k Hypothesen. Wiss. Zeitschr. d. Techn. Hochsch. Karl-Marx-Stadt, 23, 361-367.

[30] EGER, K.-H. (1984). A direct method of computation of the characteristics for SPRTs in the case of discrete random variables. in 'Sequential Methods of Statistics', Banach Center Publications, volume 16, Warsaw, 73-84.

[31] EGER, K.-H., FLEISCHER, W. (1981). Sequentielle Prüfpläne für Poissonverteilte Fehlermerkmale. in "Qualitätsanalyse 1981", Verlag Kammer der Technik Frankfurt/Oder.

[32] EWAN, W.D., KEMP, K.W. (1960). Sampling inspection of continuous processes with no autocorrelation between successive results. Biometrika 47, 363-380.

[33] FRANZ, J., WINKLER, W. (1976). Über stoppzeiten bei statistischen Problemen für homogene Prozesse mit unabhängigen Zuwächsen. Mathematische Nachrichten 70, 37-53.

[34] GHOSH, B.K. (1969). Moments of the distribution of sample size in a SPRT. JASA 64, 1560-1575.

[35] GHOSH, B.K. (1970). Sequential tests of statistical hypotheses. Addison-Wesley, Reading.

[36] GIRSHICK, M.A. (1946). Contributions to the theory of sequential analysis II + III. Ann. Math. Statist. 17, 282-298.

[37] GOVINDARAJULU, Z.G. (1981). The sequential statistical analysis of hypotheses-testing, point and interval estimation and decision theory. American Sciences Press (Columbia Ohio).

[38] HALD, A., MØLLER,U. (1976). Multiple sampling plans of given strength for the Poisson and binomial distribution. Preprint No. 11, Inst. Math. Statist.Univ. Copenhagen.

[39] HALL, W.J. (1970). On Wald's equations in continuous time. J. appl. Prob. 7, 59-68.

[40] HECKENDORFF, H. (1982). Grundlagen der sequentiellen Statistik. Teubner-Texte zur Mathematik, Band 45, BSB B. G. Teubner Verlagsgesellschaft, Leipzig.

[41] HOEFFDING, W. (1953). A lower bound for the average sample number of a sequential test. Ann. Math. Statist. 24, 127-130.

[42] HOEFFDING, W. (1960). Lower bounds for the expected sample size and the average risk of a sequential procedure. Ann. Math. Statist. 31, 352-368.

[43] HOEL, D.G. (1970). On the monotonicity of the OC of an SPRT. Ann. Math. Statist. 41, 310-314.

[44] HOEL, D.G., WEISS, G.H., SIMON, R. (1976). Sequential tests for composite hypotheses with two binomial populations. JRSS, Ser. B, 38, 302-308.

[45] IRLE, A., SCHMITZ, N. (1984). On the optimality of the SPRT for processes with continuous time parameter. Math. Operationsforsch. Statist., Ser. Statistics, 15, 91-104.

[46] JACKSON, J.E. (1960). Bibliography on sequential analysis. JASA 55, 561-580.

[47] JOHNSON, N.L. (1959). A proof of Wald's theorem on cumulative sums. Ann. Math. Statist. 30, 1245-1247.

[48] JOHNSON, N.L. (1961). Sequential analysis: a survey. JRSS, Ser. A, 124, 372-411.

[49] JONES, H.L. (1952). Formulas for the group sequential sampling of attributes. Ann. Math. Statist. 23, 72-87.

[50] KEMP, K.W. (1958). Formulae for calculating the operating characteristics and the averge sample number of some sequential tests. JRSS, Ser. B, 20, 379-386.

[51] KIEFER, J., WEISS, L. (1957). Some properties of generalized sequential probability ratio tests. Ann. Math. Statist. 28, 57-74.

[52] LECHNER, J.A. (1964). Optimality and the OC curve for the Wald SPRT. JASA 59, 464-468.

[53] LEHMANN, E.L. (1959). Testing Statistical Hypotheses. Wiley, New York.

[54] LORDEN, G. (1980). Structure of sequential tests minimizing an expected sample size. Zeitschr. Wahrscheinlichkeitstheorie verw. Geb. 51, 291-302.

[55] MALÝ, V. (1960). Sequenzpropleme mit mehreren Entscheidungen und Sequenzschätzung. Biometrische Zeitschr. 2, 45-64.

[56] MALÝ, V. (1966). Zu Fragen des Sequenzvergleichs von zwei relativen Häufigkeiten beim Paarversuch. Biometrische Zeitschr. 8, 162-178.

[57] MANLY, B.F. (1970). On the distribution of the DSN of certain sequential tests. Biometrika 57, 367-376.

[58] MATTHES, T.K. (1963). On the optimality of sequential probability ratio tests. Ann. Math. Statist. 34, 18-21.

[59] MC NEIL, D.R. (1966). Grouping corrections for sequential tests. Austral. J. Statist. 8, 9-21.

[60] PAGE, E.S. (1954). An improvement to Wald's approximation for some properties of sequential tests. JRSS, Ser. B, 16, 136-139.

[61] PAGE, E.S. (1954). Continuous inspection schemes. Biometrika 41, 100-115.

[62] PAGE, E.S. (1957). On problems in which a change in a parameter occurs at an unknown time point. Biometrika 44, 248-252.

[63] PAULSON, E. (1963). A sequential decision procedure for choosing one of k hypotheses concerning the unknown mean of a normal distribution. Ann. Math. Statist. 34, 549-555.

[64] PÓLYA, G. (1948). Exact formulas in the sequential analysis of attributes. University of California Publications in Mathematics, New Series, 1, 229-239.

[65] RAY, W.D. (1957). A proof that the sequential probability ratio test of the general linear hypothesis terminates with probability unity. Ann. Math. Statist. 28, 521-523.

[66] SAVAGE, L.J. (1957). When different pairs of hypotheses have the same family of likelihood-ratio test regions. Ann. Math. Statist. 28, 1028-1032.

[67] SCHMITZ, N. (1973). Zur Optimalität des Likelihoodquotienten-Sequenztests. Transactions Sixth Prague Conference 1971, 781-791.

[68] SCHMITZ, N. (1982). Sequential probability ratio tests for stochastic processes - a review note. Arbeitspapier S 23 des Instituts für Mathematische Statistik de Universität Münster.

[69] SIMON, R., WEISS, G.H., HOEL, D.G. (1975). Sequential analysis of binomial clinical trials. Biometrika 62, 195-200.

[70] SIMONS, G. (1967). A sequential three hypothesis test for determining the mean of a normal population with known variance. Ann. Math. Statist. 38, 1365-1375.

[71] SIMONS, G. (1967). Lower bounds for average sample number of sequential multihypothesis test. Ann. Math. Statist. 38, 1343-1364.

[72] SOBEL, M., WALD, A. (1949). A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. Ann. Math. Statist. 20, 502-522.

[73] ŠIRJAEV, A.N. (1976). Statisticeskij posledovatel'nyi analiz. Nauka, Moskva.

[74] STEIN, C, (1946). A note on cumulative sums. Ann. Math. Statist. 17, 488-499.

[75] TRYBULA, S. (1968). Sequential estimation in processes with independent increments. Dissertationes Mathematicae LX, Warsaw.

[76] WALD, A. (1944). On cumulative sums of random variables. Ann. Math. Statist. 15, 283-296.

[77] WALD, A. (1947). Sequential analysis. Wiley, New York.

[78] WALD, A., WOLFOWITZ, J. (1950). Optimum character of the sequential probability ratio test. Ann. Math. Statist. 19, 326-339.

[79] WALKER, A.M. (1950). Note on sequential sampling formulae for a binomial population. JRSS, Ser. B, 12, 201-207.

[80] WEISS, L. (1953). Testing one simple hypothesis against another. Ann. Math. Statist. 24, 273-281.

[81] WIJSMAN, R.A. (1971). Exponentially bounded stopping time of sequential probability ratio test for composite hypotheses. Ann. Math. Statist. 42, 1859-1869.

[82] WIJSMAN, R.A. (1975). Exponentially bounded stopping time of the sequential t-test. Ann. Statist. 3, 1006-1010.

[83] WOLFOWITZ, J. (1947). Efficiency of sequential estimates and Wald's equation for sequential processes. Ann. Math. Statist. 18, 215-230

[84] WOLFOWITZ, J. (1966). Remark on tne optimum character of the sequential probability ratio test. Ann. Math. Statist. 37, 726-727.

ALGEBRAIC AND DIFFERENTIAL TOPOLOGY - GLOBAL DIFFERENTIAL GEOMETRY
Edited by George M. Rassias

This Teubner-Text contains a series of specially invited research and expository papers written by eminent scientists from the international community on the occasion of the 90th anniversary of Marston Morse's birth.
This commemorative volume provides an insight and analysis on various research problems and recent new theories in the fastest developing and very interesting fields of Algebraic and Differential Topology, Global Differential Geometry and Related Topics such as the Global Analysis of Differential Equations (Ordinary and Partial) and of the Calculus of Variations.
Bd. 70, 348 Seiten, 1984, M 36,--

H. HECKENDORFF
Grundlagen der sequentiellen Statistik

In diesem Text wird ein systematischer Zugang zur Darstellung der Grundlagen der Theorie des sequentiellen statistischen Experiments erarbeitet. Dieser Zugang beinhaltet die für die Anwendung stets gleichermaßen bedeutsamen Aspekte der Festlegung der Strategie der Versuchsplanung, des Umfangs des Experiments sowie der auf den entstehenden Stichprobenräumen basierenden Entscheidungsregeln. Dabei wird der gerade im sequentiellen Fall besonders relevanten Nutzung des Prinzips der Suffizienz besondere Aufmerksamkeit gewidmet. Ferner wird die sequentielle Bayessche Entscheidungstheorie in gewisser Gründlichkeit eingeordnet. Die Theorie der sequentiellen Versuchsplanung wird für die Aufgabe des zweiarmigen Banditen und für das Regressionsproblem spezifiziert.
Bd. 45, 166 Seiten, 1982, M 18,--

V. G. KORNEEV/U. LANGER
Approximate Solution of Plastic Flow Theory Problems

The book is devoted to several aspects of the numerical solution of boundary value problems in the flow theory of plasticity and may be used as a text-book for university students and postgraduates specializing in numerical analysis and mechanics. The main boundary value problems, including problems which are ill-posed, and various methods of regularizing them, are considered together with some preliminary results obtained from the investigation of solvability in Sobolev spaces. The numerical schemes considered are based on the finite element method, the incremental loading method and some iterative methods for the solution of plastic flow theory problems and the systems of nonlinear algebraic equations arising in the process of their numerical solution.
Bd. 69, 252 Seiten, 1984, N. 26,--

S. L. KRUSCHKAL/R. KÜHNAU
Quasikonforme Abbildungen - neue Methoden und Anwendungen

Es werden einige neuere Methoden und Entwicklungen, Ergebnisse und Anwendungen der Theorie der quasikonformen Abbildungen dargestellt. Zur Behandlung kommen Extremal- und Wertannahmeprobleme bei ortsabhängiger Dilatationsbeschränkung, u. a. auch mit der Methode der biholomorph invarianten Metriken, allgemeine Darstellungssätze für Extremalfunktionen, einfache asymptotische Abschätzungen. Es werden Anwendungen auf Extremalprobleme gegeben, z. B. in der ebenen Elektrostatik bei inhomogenen Medien. Der Zusammenhang zwischen Teichmüllerschen Räumen und schlichten Funktionen mit einer quasikonformen Fortsetzung wird ebenfalls beleuchtet.
Bd. 54, 172 Seiten, 1983, M 17,50

B. LISEK/J. HOCHSCHILD

Sequentielle Zuverlässigkeitsprüfung

Die statistische Zuverlässigkeitsprüfung ist ein Gebiet mit großer praktischer Bedeutung. Für den Fall exponentialverteilten Ausfallabstands werden alle damit zusammenhängenden Fragen ausführlich diskutiert, wobei sequentielle Prüfungen im Mittelpunkt stehen. Neue Ergebnisse betreffen die Möglichkeit extrem kurzer Stutzung sequentieller Prüfungen. Größter Wert wird auf Modelldiskussionen und Interpretationen gelegt. So ist die Darstellung auch für Nichtmathematiker verständlich. Es ist eine große Zahl von Tabellen aufgenommen worden, die für den Zuverlässigkeitspraktiker nützlich sein werden.
Bd. 53, 152 Seiten, 1983, M 16,--

W. NÄTHER

Effective Observation of Random Fields

The book deals with designing methods for linear estimation of the trend and for linear prediction of random processes and fields with known covariance function. Especially, effective observation methods for the least-squares estimator and the best linear predictor are considered (chapters 4 - 8). In these chapters the main goal consists in demonstrating to what an extent classical convex designing methods (chapter 3) can be modified to yield useful results in the process case. Besides exact, iterative and asymptotic procedures also approximative methods are proposed. The remaining part is devoted to the problems of the optimal choice of an observation region, of weakening the assumption of a known covariance function and deals with designing methods using Fisher information (chapters 9 - 11).
Bd. 72, 184 Seiten, 1985, M 18,--

L. PARTZSCH

Vorlesungen zum eindimensionalen Wienerschen Prozeß

Das vorliegende Buch entstand aus einer einführenden Vorlesung über den eindimensionalen Wienerschen Prozeß. Sein Hauptanliegen ist eine ausführliche Darstellung der klassischen Grundlagen: Konstruktiver Existenzbeweis, Invarianzeigenschaften des Wienerschen Maßes, Bestimmung der Verteilung ausgewählter Funktionale, lokale und globale Eigenschaften der Realisierungen. Anschließend werden zwei speziellere Probleme behandelt, nämlich die Skorochod-Darstellung und das Strassensche Gesetz des iterierten Logarithmus (hierbei spielen eine Cameron-Martin-Formel und Exponentialabschätzungen für das Wienersche Maß eine wesentliche Rolle). Vom Leser werden nur einfache Kenntnisse über stochastische Prozesse vorausgesetzt.
Bd. 66, 112 Seiten, 1984, M 12,--

J. PILZ

Bayesian Estimation and Experimental Design in Linear Regression Models

The book deals with estimation and experimental design for linear regression models in the presence of prior knowledge about the model parameters; the approach is Bayesian throughout. The first part starts with a formulation of the compound problem of estimation and design within the framework of Bayesian decision theory, and then proceeds with the specification of a prior probability distribution for the model parameters. The second part is concerned with the natural-conjugate Bayes estimator under normally, independently and identically distributed errors and quadratic loss, and studies the robustness of optimality under a change of the model assumptions; furthermore, some relations with other important alternative to the least squares estimator are established. The third part is devoted to the experimental design problem for the Bayes estimator and deals

extensivly with the construction of both optimal approximate and optimal exact designs.
Bd. 55, 216 Seiten, 1983, M 20,--

RECENT TRENDS IN MATHEMATICS, REINHARDSBRUNN 1982
Editors: H. Kurke/J. Mecke/H. Triebel/R. Thiele

The conference "Recent Trends in Mathematics" was held at Reinhardsbrunn, GDR, from October 11 to October 13, 1982. It was organized by the BSB B. G. Teubner Verlagsgesellschaft, Leipzig, and was attended by mathematicians from Austria, Canada, Czechoslovakia, Finland, France, FRG, Great Britain, Italy, Japan, Poland, Poumania, Sweden, U.S.A., USSR, and GDR. This fiftieth Teubner-Text contains 31 lectures given at this conference.
Bd. 50, 336 Seiten, 1983, M 30,--

H.-U. SCHWARZ
Banach Lattices and Operators

The book contains an introduction to the theory of vector lattices, Banach lattices, and bounded operators in Banach lattices. The theory of vector lattices is developed as far as it is needed for further investigation of Banach lattices. In the second part, which is concerned with Banach lattices, the main emphasis lies on the presentation of various classes of Banach lattices, as order complete spaces, KB-spaces, and others. All lattice theoretic notions are discussed in classical Banach spaces and in Köthe function spaces. The third part deals with basic order properties of linear operators, and some special classes of operators. The results are used to study the structure of Banach lattices.
Bd. 71, 208 Seiten, 1984, M 19,50

STOCHASTIC GEOMETRY, GEOMETRIC STATISTICS, STEREOLOGY
Editors: R. Ambartzumian/W. Weil

This volume contains contributions to the conference "Stochastic Geometry, Geometric Statistics, Stereology" in Oberwolfach, 1983, as well as a number of further papers in this field. The topics cover various parts of Stochastic Geometry: random sets, integral geometry, point processes, statistical geometry, random coverings, random mosaics, stereology.
Bd. 65, 268 Seiten, 1984, M 27,50

F. TRÖLTZSCH
Optimality Conditions for Parabolic Control Problems and Applications

This text is concerned mainly with necessary optimality conditions for optimal control problems governed by non-linear parabolic differential equations, where the control may occur on the boundary or within the domain under consideration. Thereby constraints on the state as well as on the control may be given. The theory is developed up to the proof of bang-bang properties for optimal controls. The problems are treated by converting them into a control problem for a Hammerstein integral equation in a Banach space and by applying Lagrange multiplier rules for abstract mathematical programs. Therefore, the teyt also contains a chapter on the optimal control of Hammerstein integral equations and an introduction to the Kuhn-Tucker theory for non-linear mathematical programming problems in Banach spaces.
Bd. 62, 164 Seiten, 1984, M 17,--

Diese Reihe wurde geschaffen, um eine schnellere Verö
mathematischer Forschungsergebnisse und eine weitere
mathematischen Spezialvorlesungen zu erreichen. TEUB
in deutsch, englisch, russisch oder französisch ersc
alität der Reihe zu erhalten, werden die TEUBNER-TE
druck hergestellt, da so die geringeren drucktechnis
eine raschere Herstellung ermöglichen. Autoren von TE
liefern an den Verlag ein reproduktionsfähiges Manusk
Auskünfte darüber erhalten die Autoren vom Verlag.

This series has been initiated with a view to quicker publication of the results of mathematical research-work and a widespread circulation of special lectures on mathematics. TEUBNER-TEXTE will be published in German, English, Russian or French. In order to keep this series constantly up to date and to assure a quick distribution, the copies of these texts are produced by a photographic process (small-offset printing) because its technical simplicity is ideally suited to this type of publication. Authors supply the publishers with a manuscript ready for reproduction in accordance with the latter's instructions.

Cette série des textes a été créée pour obtenir une publication plus rapide de résultats de recherches mathématiques et de conférences sur des problèmes mathématiques spéciaux. Les TEUBNER-TEXTE seront publiés en langues allemande, anglaise, russe ou française. L'actualité des TEUBNER-TEXTE est assurée par un procédé photographique (impression offset). Les auteurs des TEUBNER-TEXTE sont priés de fournir à notre maison d'édition un manuscrit prêt a être reproduit. Des renseignements plus précis sur la forme du manuscrit leur sont donnés par notre maison.

Эта серия была создана для обеспечения более быстрого публикования результатов математических исследований и более широкого распространения математических лекций на специяльные темы. Издания серии ТОЙБНЕР-ТЕКСТЕ будут публиковаться на немецком, английском, русском или французском языках. Для обеспечения актуальности серии, ее издания будут изготовляться фотомеханическом способом. Таким образом, более скромные требования к полиграфицескому оформлению обеспечат более быстрое появление в свет. Авторы изданий серии ТОЙБНЕР-ТЕКСТЕ будут предоставлять издательству рукописи, удовлетворящие требованиям фотомеханического печатания. Более подробные сведения авторы получат от издательства.

BSB B. G. Teubner Verlagsgesellschaft, Leipzig
DDR-701 Leipzig, Postfach 930

TEUBNER-TEXTE zur Mathematik

ALGEBRAIC AND DIFFERENTIAL TOPOLOGY - GLOBAL DIFFERENTIAL GEOMETRY
Editor: G. Rassias

GLOBAL ANALYSIS - ANALYSIS ON MANIFOLDS
Editor: T. Rassias

GRAPHS AND OTHER COMBINATORIAL TOPICS
Editor: Fiedler

HECHENDORFF
Grundlagen der sequentiellen Statistik

JUNEK
Locally Convex Spaces and Operator Ideals

KORNEEV/LANGER
Approximate Solution of Plastic Flow Theory Problems

KRUSCHKAL/KÜHNAU
Quasikonforme Abbildungen

LINDE
Infinitely Divisible and Stable Measures on Banach Spaces

LISEK/HOCHSCHILD
Sequentielle Zuverlässigkeitsprüfung

OBERGUGGENBERGER
Der Graphensatz in lokalkonvexen topologischen Vektorräumen

PARTZSCH
Vorlesungen zum eindimensionalen Wienerschen Prozeß

PILZ
Bayesian Estimation and Experimental Design in Linear Regression Models

RECENT TRENDS IN MATHEMATICS, REINHARDSBRUNN 1982
Editors: Kurke/Mecke/Triebel/Thiele

RENELT
Quasikonforme Abbildungen und elliptische Systeme

SEMADENI/WIWEGER
Einführung in die Theorie der Kategorien und Funktoren

STOCHASTIC GEOMETRY, GEOMETRIC STATISTICS, STEREOLOGY
Editors: Ambartzumian/Weil

TRÖLTZSCH
Optimality Conditions for Parabolic Control Problems and Applications

TUTSCHKE
Partielle Differentialgleichungen. Klassische funktionalanalytische und komplexe Methoden

ZEIDLER
Vorlesungen über nichtlineare Funktionalanalysis

ZINK
Cartiertheorie kommutativer formaler Gruppen

In preparation:

DIFFERENTIAL TOPOLOGY
Editor: G. Rassias

ENGEL/GRONAU
Sperner Theory in Partially Ordered Sets

GRAPHS, HYPERGRAPHS AND APPLICATIONS
Editor: Sachs

GRIEPENTROG/MÄRZ
Differential Algebraic Equations and Their Numerical Treatment

HUBER-DYSON
A Work Book on Gödel's Theorems

KAČUR
Method of Rothe in Evolution Equations

KOSHELEV
Regularity of the Solution of Elliptic Equations and Systems of Equations

KUFNER/SÄNDIG
Some Applications of Weighted Sobolev Spaces

MAREK/ŽITNÝ
Matrix Analysis for Applied Sciences · Vol. 2

MELTZER
The Structure of Indecomposable Modules

NÄTHER
Effective Observation of Random Fields

T. RASSIAS
Foundations of Global Nonlinear Analysis

WALLISCH/HERMANN
Schießverfahren zur Lösung von Rand- und Eigenwertaufgaben